



Instrumentation pour le radiotélescope de 30 mètres de l'IRAM et observations d'atmosphères planétaires

Gabriel Paubert

► To cite this version:

Gabriel Paubert. Instrumentation pour le radiotélescope de 30 mètres de l'IRAM et observations d'atmosphères planétaires. Instrumentation et méthodes pour l'astrophysique [astro-ph.IM]. Université Joseph-Fourier - Grenoble I, 1992. Français. NNT: . tel-00688090

HAL Id: tel-00688090

<https://theses.hal.science/tel-00688090>

Submitted on 16 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE
PRESENTEE PAR
Gabriel PAUBERT

POUR OBTENIR LE TITRE DE DOCTEUR
DE L'UNIVERSITE JOSEPH FOURIER - GRENOBLE 1
(ARRETES MINISTERIELS DU 5 JUILLET 1984 ET
DU 23 NOVEMBRE 1988)
SPECIALITE PHYSIQUE

**Instrumentation pour le radiotélescope de 30 mètres de l'IRAM et
observations d'atmosphères planétaires**

DATE DE SOUTENANCE : 30 Octobre 1992

COMPOSITION DU JURY :

Guy PELLETIER	Président
Françoise COMBES	Rapporteur
Pierre ENCRENAZ	Rapporteur
Bernard LAZAREFF	Examineur
André MARTEN	Examineur

THESE PREPAREE AU SEIN DU
LABORATOIRE D'ASTROPHYSIQUE DE GRENOBLE

Remerciements

Ce travail n'aurait jamais pu être mené à bien sans le concours de nombreuses personnes. C'est d'autant plus vrai que ma tendance à me disperser plus que de raison m'a fait aborder des sujets très variés.

Parmi toutes ces personnes, mon directeur de thèse, Bernard Lazareff, mérite une mention spéciale. Sans ses encouragements, son dévouement et ses commentaires très indulgents sur le manuscrit, je n'aurais jamais pu mener à bien la rédaction de ce mémoire. Je dois aussi exprimer ma gratitude au précédent directeur de l'IRAM, Peter de Jonge, qui m'a proposé de continuer à travailler à Grenade après mon service de coopération et sans qui ce travail n'aurait évidemment jamais été entrepris.

Le mode de fonctionnement du télescope, avec les astronomes qui viennent y observer, fournit l'occasion de nouer de nombreux contacts et d'aborder des sujets très variés. C'est un environnement très enrichissant, qui permet au passage de se perfectionner dans la langue de Shakespeare et ses métamorphoses phonétiques, qui atteignent un point où on ne peut plus parler d'accent. Contrairement à ce que l'on pourrait croire, les plus incompréhensibles sont quelquefois les anglais de pure souche.

Quant à la ville de Grenade, entre ses monuments et quartiers historiques où le temps semble s'arrêter, son quartier étudiant très animé avec ses innombrables bars, elle réserve à l'étranger qui vient y vivre un excellent accueil, une fois passées les difficultés initiales causées par la barrière de la langue. Son climat a transformé en mauvais souvenir la grisaille de cinq hivers passés à Paris, quelquefois déprimants pour quelqu'un qui a toujours vécu dans le Sud-Ouest.

Dans les conditions parfois adverses rencontrées au télescope, il faut particulièrement remercier les opérateurs, toujours prêts à retoucher les réglages des récepteurs et à récupérer les données effacées par mégarde, et bien entendu les cuisinières dont les repas, particulièrement appréciés, constituent un moment de détente indispensable, soit que l'on se trouve confronté à des problèmes inextricables, soit que l'on ronge son frein en attendant que le temps s'améliore ou pire, que le vent veuille bien tomber.

Mes divagations scientifiques et techniques, sans compter les problèmes administratifs, m'ont conduit à travailler avec presque tout le personnel de l'IRAM Grenade, détailler dans quelles circonstances serait bien trop long, aussi me contenterai-je de donner une liste de noms dans l'ordre alphabétique: Yves Bortolotti, Walter Brunswig, Gilles Butin, Carmen Carpintero, Pierre Colom, Didier Despois, Gloria España, Esther Franzin, Pivo Garcia, Albert Greve, Hauke Hein, Christian Hôte, Hervé Houeix, Carsten Kömpe, Stella Liechti, Javier Lobato, Jean-Marc Mariani, Rainer Mauersberger, Santiago Navarro, Bernard Para, Juan Peñalver, Salvador Sanchez, Hans Steppe et Clemens Thum méritent tous ma reconnaissance.

Je dois exprimer toute ma gratitude à l'équipe du spectrographe acousto-optique qui m'a aimablement accueilli lors de mes séjours à l'Observatoire de Meudon: Patrick Dierich,

Alain Lecacheux, Micheline Pépin et Carlo Rosolen. Je ne dois pas oublier non plus Rémy Bellenger qui m'a ouvert les portes de son laboratoire pour pouvoir développer le prédifusé programmable qui joue le rôle de chef d'orchestre de l'électronique d'acquisition du spectrographe.

Il me faut bien sûr aussi remercier les astronomes avec qui j'ai eu l'occasion d'observer au radiotélescope sur des projets communs: Bruno Bézard, Dominique Bockelée-Morvan, José Cernicharo, Athéna Coustenis, Jacques Crovisier, Thérèse Encrenaz, Emmanuel Lellouch, André Marten et Jan Rosenqvist.

J'exprime toute ma gratitude au président du jury Guy Pelletier, et aux rapporteurs Françoise Combes et Pierre Encrenaz qui ont accepté de lire ce mémoire en un temps très court.

Enfin, il me faut bien remercier la patience du lecteur qui lira ma prose. Je lui demande aussi de bien vouloir me pardonner les nombreuses coquilles et fautes de frappe qui se sont glissées dans le manuscrit; chaque fois que je le relis après l'avoir corrigé, j'en retrouve à peu près autant.

Résumé

Les exigences en matière de stabilité et de linéarité des instruments utilisés en radioastronomie sont particulièrement sévères. Ce mémoire présente la conception et la réalisation d'un détecteur continuum adapté à ce cahier des charges. Une étude théorique démontre que le détecteur à diode Schottky ne répond pas à ces besoins. L'instrument finalement réalisé, construit à partir d'une diode tunnel, possède un bruit proche de la limite physique fondamentale. Il reste cependant très simple.

Le chapitre suivant est consacré à l'étude et à la réalisation de l'électronique d'acquisition d'un spectrographe acousto-optique à large bande et à l'étude de ses caractéristiques dans des conditions d'utilisation réelles. Il inclut la description d'une méthode simple qui permet de s'affranchir des erreurs de linéarité différentielle du convertisseur analogique-numérique. La validation des caractéristiques du spectrographe s'appuie sur la comparaison avec deux batteries de filtres identiques.

Sur le plan scientifique, il présente l'observation d'objets du système solaire entourés d'une atmosphère et l'interprétation des résultats dans le cas de Titan.

Abstract

Stability and linearity requirements of instruments used in radioastronomy are especially severe. This thesis describes on one hand the design and construction of a continuum detector fulfilling these requirements. A theoretical study demonstrates that the Schottky diode detector can not provide satisfactory characteristics. The deetector built here, built around tunnel diodes, exhibits a noise close to the limits imposed by fundamental physics. It has nonetheless been kept simple.

The following chapter describes the design and construction of the acquisition electronics for a wideband acousto-optic spectrometer and studies its characteristics under realistic operating conditions. It includes the description of a simple method which allows to get rid of the differential linearity errors of analog-to-digital converters. The validation of the spectrograph performance is based on the comparison with two other, identical, filterbank spectrometers.

On the scientific side, the reports presents the observations of several objects of the solar system surrounded by an atmosphere and the interpretation of the data in the case of Titan.

Table des Matières

1	Introduction	11
2	Radioastronomie	15
2.1	Architecture d'un radiotélescope	15
2.1.1	L'antenne	15
2.1.2	Le ou les récepteurs.	16
2.1.3	Spectrographes	18
2.2	Méthodes d'observation	20
2.2.1	Modulation par commutation de lobe	21
2.2.2	Modulation par commutation de fréquence	21
2.2.3	Détection synchrone	22
2.2.4	Etalonnage	22
2.2.5	Linéarité	24
2.2.6	Dynamique	26
2.2.7	Stabilité	27
2.2.8	Variances d'Allan	27
3	Continuum	29
3.1	Motivation	29
3.1.1	Observations en continuum	29
3.1.2	Pointage	29
3.1.3	Focalisation	30
3.1.4	Efficacité de l'antenne et opacité de l'atmosphère	30
3.1.5	Coefficient de couplage du lobe principal	31
3.1.6	Mesures techniques	31
3.2	Caractéristiques requises	31
3.3	Mesures de puissance en radiofréquence	32
3.4	Terminologie des détecteurs à diode	33
3.5	Le détecteur à diode Schottky	34
3.5.1	Réponse en puissance du détecteur	35
3.5.2	Fluctuation du signal détecté	37
3.5.3	Puissance minimale utilisable	40
3.5.4	Conséquences	41
3.6	Diodes tunnel	41
3.7	Contrôle thermique	43
3.8	Amplification	44
3.9	Réalisation pratique	46
3.10	Acquisition	48

3.10.1	Considérations générales	48
3.10.2	Conversion tension-fréquence	48
3.10.3	Schéma théorique	50
3.10.4	Réalisation pratique	53
3.11	Réglages	54
3.12	Maintenance	54
3.13	Mesure des caractéristiques en laboratoire	54
3.13.1	Niveau de bruit et autre paramètres électroniques	54
3.13.2	Linéarité en puissance	55
3.13.3	Ondulation de la réponse en fréquence	57
3.13.4	Taux d'onde stationnaire	57
3.14	Mesures au télescope	58
3.14.1	Variances d'Allan	59
3.14.2	Linéarité en puissance	61
3.15	Exemples d'observations	62
3.15.1	Pointage	62
3.15.2	Détermination du flux d'un quasar	62
3.16	Mesure de la stabilité des récepteurs	62
3.17	Holographie	64
3.18	Remarques	64
3.19	Conclusion et perspectives	65
4	Spectrographe acousto-optique	67
4.1	Introduction	67
4.2	L'interaction acousto-optique	67
4.2.1	Point de vue de l'optique classique	67
4.2.2	Point de vue corpusculaire	68
4.2.3	Point de vue électrodynamique dans la matière	69
4.3	Application à la spectroscopie	72
4.3.1	Matériaux utilisés	73
4.4	Le spectrographe installé à l'IRAM	74
4.4.1	Architecture générale	74
4.4.2	La cellule de Bragg	74
4.4.3	Source de lumière	74
4.4.4	Optique et photodétecteur	76
4.4.5	Electronique d'acquisition et de contrôle	76
4.4.6	Linéarité	78
4.4.7	Processeur de fréquence intermédiaire	81
4.4.8	Processeur local	82
4.4.9	Interface avec l'ordinateur de contrôle du télescope	82
4.5	Caractéristiques mesurées	82
4.5.1	Comparaison avec les batteries de filtres	82
4.5.2	Filtre équivalent	83
4.5.3	Variance d'Allan	84
4.5.4	Ondulation de la bande	87
4.5.5	Dynamique	88
4.5.6	Etalonnage en fréquence	89
4.5.7	Dérives à long terme du spectrographe	90

4.6	Conclusion	92
5	Système solaire	95
5.1	Introduction	95
5.2	Titan	95
5.2.1	Introduction	95
5.2.2	Transfert de rayonnement	97
5.2.3	Observations	97
5.2.4	Acide cyanhydrique	98
5.2.5	Monoxyde de carbone	98
5.3	Autres observations	102
5.3.1	Io	102
5.3.2	Mars	103
5.3.3	Comètes	103
A	Measurements of receiver power stability	105
A.1	Introduction	105
A.2	Technical details	105
A.3	Results	106
A.4	Conclusions	106
B	Technical documentation of the AOS acquisition system	113
B.1	Overview	115
B.2	Architecture	115
B.3	Analog part	117
B.3.1	CCD board	117
B.3.2	Analog board	117
B.3.3	Adjustments	119
B.4	Digital board	121
B.5	Sequencer	121
B.5.1	Pixel level sequencing	123
B.5.2	Frame level sequencing	123
B.5.3	Sequencer implementation	126
B.5.4	Modifying the sequencer	127
B.6	Bus interface	128
B.7	Power-up and hardware reset	128
B.8	Possible upgrades	129
B.9	LCA layout	130
B.10	Programming	130
B.10.1	Address map	130
B.10.2	Operation mode programming	133
B.10.3	Status words	134
B.10.4	Interrupts	135
B.10.5	Programming hints	136
B.11	Synchronization between AOS	136
B.12	Power requirements	139
B.13	Known problems	139

C	On the reduction of IRAM-30M Io data	141
C.1	Introduction	141
C.2	Notations, Definitions and Hypothesis	141
C.3	The resampling operation	143
C.4	Filter shape	143
C.5	Line width	143
C.6	Autocorrelation bandwidth	144
C.7	Integrated line intensity	145
C.8	Correlation between channels	145
C.9	Other sources of error	146
C.10	Conclusion	146

Chapitre 1

Introduction

Le radiotélescope de 30 mètres de l'IRAM [20], situé à 2870 mètres d'altitude près du Pico Veleta dans la Sierra Nevada, est un instrument unique par ses caractéristiques. La taille de l'antenne, combinée à la précision de sa surface, lui confère une sensibilité et une résolution inégalées dans la gamme de longueur d'onde pour laquelle il a été conçu: entre 1 et 3 mm, plus précisément entre 80 et 350 GHz, puisque les radioastronomes expriment généralement les valeurs exactes en fréquence et les valeurs approximatives en longueur d'onde.

Cette partie du spectre électromagnétique, pour laquelle le premier grand télescope fut l'antenne de 12 mètres de Kitt Peak en Arizona, présente au moins deux caractéristiques intéressantes.

Elle correspond tout d'abord aux premières raies de transition de rotation de radicaux, d'ions et de molécules simples et par conséquent relativement abondants, comme CN, HCO^+ , HCN, sans oublier le monoxyde de carbone CO, de loin le corps le plus étudié. Les raies de ces composés linéaires correspondent à de petites corrections près aux harmoniques d'une fréquence fondamentale; pour des géométries moins simples, lorsque le tenseur d'inertie possède 2 ou 3 valeurs propres distinctes, la loi qui gouverne la distribution des raies devient nettement plus complexe. L'observation de composés polyatomiques relativement lourds est bien sûr aussi possible, et fournit des indications capitales sur la chimie des enveloppes d'étoiles et du milieu interstellaire. Il faut cependant souligner qu'on ne peut détecter que les molécules qui possèdent un moment dipolaire non nul, ce qui exclut la catégorie importante des molécules diatomiques homonucléaires, et quelques autres comme le méthane, l'acétylène (alors que l'on observe souvent des nitriles de chaînes acétyléniques comme HC_5N), etc. Cette règle souffre quelques exceptions, dont la plus connue est l'oxygène, de toutes façons très difficile à observer à travers l'atmosphère terrestre.

L'autre point important est une simple considération énergétique: les raies de rotation sont les plus faciles à exciter et on peut les observer dans des corps très froids, comme les nuages moléculaires sombres et denses où la température est comprise entre 10 et 20 K. Ces objets, qui sont complètement opaques dans les gammes visible et infrarouge, sont le siège de phénomènes d'effondrement gravitationnel qui sont à l'origine de la formation des étoiles: l'un des problèmes fondamentaux de l'astronomie. Remarquons ici que la basse température de ces objets implique que la dispersion en vitesse ne sera que de quelques dizaines de mètres par seconde, ce qui exige une très haute résolution relative, de l'ordre de 10^7 , si l'on veut résoudre les raies. On y observe des corps dont les spectres sont difficiles à déterminer du fait de leur réactivité, et les observations astronomiques donnent quelquefois des valeurs précises de leurs paramètres spectroscopiques avant même d'être confirmé en laboratoire [21]. Dans d'autres cas, et très souvent ces dernières années grâce au radiotélescope de 30 mètres,

ces corps sont observés dès que les constantes permettant de prédire leurs spectres ont été déterminées [16, 22].

Une partie importante du temps d'observation du radiotélescope est aussi consacré à l'étude d'objets extragalactiques. Il s'agit essentiellement d'observation des nuages moléculaires par l'intermédiaire du monoxyde de carbone, traceur du gaz dense, ce qui ramène là encore au sujet de la formation des étoiles. Les raies que l'on observe sont en général très larges car l'on voit la dispersion en vitesse de tous les nuages contenus dans le lobe du télescope. Cette dispersion atteint souvent plusieurs centaines de kilomètres par seconde, ce qui présente des problèmes de couverture instantanée en fréquence, qui est limitée à une fraction de pourcent. Ces observations fournissent des informations sur la dynamique et l'évolution de ces systèmes extrêmement complexes que sont les galaxies.

Dans un tout autre domaine, le système solaire, les observations effectuées ces dernières années à l'IRAM sont extrêmement variées bien qu'elles aient un point commun: elles ont toujours porté sur les enveloppes gazeuses des objets en question, que celles-ci soient denses et stables comme dans le cas de Titan ou Neptune [23], ou ténues et éphémères comme dans les comètes. Quant aux objectifs, ils vont de la distribution des composants mineurs de l'atmosphère de Titan [2, 3] à la détermination de la circulation dans la moyenne atmosphère de Mars [5] et de Vénus, objets que l'on peut résoudre et dans lesquels on arrive à mesurer la vitesse des vents à quelques mètres par seconde près. Le profil des raies dépend énormément de l'objet étudié: entre une gaussienne de quelques centaines de kilohertz de largeur pour le cas des faibles pressions (Io, comètes) où seule la vitesse entre en compte et une Lorentzienne de quelques centaines de mégahertz (le monoxyde de carbone sur Titan) pour des pressions de l'ordre de l'atmosphère où l'élargissement collisionnel domine. On rencontre aussi des cas de superposition d'un cœur de raie étroit venant des couches supérieures de l'atmosphère avec une raie plus large formée plus profondément en toile de fond (Vénus).

La liste qui précède n'a aucune prétention d'exhaustivité mais certaines observations de caractère absolument unique effectuées avec l'antenne de 30 mètres méritent d'être mentionnées, comme l'observation d'un maser SiO lors de son occultation par la Lune qui donne une résolution d'environ 20 millisecondes d'arc (non encore publié). Le nombre d'occasions d'effectuer une observation intéressante par cette méthode est bien sûr extrêmement réduit.

L'un des points forts du radiotélescope est indiscutablement la possibilité d'utiliser plusieurs récepteurs simultanément, jusqu'à trois à l'heure actuelle et peut-être quatre dans le futur. Cela requiert évidemment la présence de suffisamment de spectrographes pour pouvoir analyser les signaux issus de tous les récepteurs. Le chapitre 4 est consacré à l'étude et à la réalisation de la partie électronique d'un spectrographe acousto-optique ainsi qu'à l'évaluation de ses performances dans le cadre de l'utilisation au télescope. Il s'agit là d'un analyseur à large bande (500 MHz) et de résolution limitée (1.1 MHz), donc plus adapté aux observations extragalactiques et d'atmosphères denses.

Il a été jusqu'ici uniquement question d'observations de raies spectrales; de nombreuses sources émettent aussi un rayonnement continu, ce qui correspond à des mécanismes d'émission et à un type d'observations totalement différents. Certaines de ces sources, tout particulièrement les quasars, exhibent d'importantes variations de « luminosité » dans le temps. Un programme de suivi de ces sources est en cours [24], qui récupère une partie des données d'une activité technique indispensable au radiotélescope: la mesure et la correction des paramètres de pointage de l'antenne.

En effet, et contrairement à ce qui se passe en astronomie optique, on ne peut pas avec un radiotélescope assurer que l'on vise la direction souhaitée dans le ciel en se référant à un objet guide situé dans le champ de visée. Comme l'antenne souffre d'inévitables petites

imperfections mécaniques, par exemple l'axe d'azimuth n'est pas parfaitement vertical, on détermine un modèle de ces imperfections en observant des sources de référence. On constate aussi que les paramètres de ce modèle dérivent pour des raisons plus ou moins bien connues: effets thermiques, vents quelquefois très violents, tremblements de terre, etc. Quelques heures chaque semaine doivent être consacrées à la mise à jour de leurs valeurs, c'est ce qu'on appelle le modèle de pointage de l'antenne. La plupart de ces observations sont effectuées sur des sources qui émettent un continuum, pour lesquelles on mesure la puissance reçue dans toute la bande passante du récepteur.

Le chapitre 3 est consacré à l'étude et à la réalisation d'un détecteur continuum de très hautes performances. L'un des buts essentiels de ce projet était d'améliorer la sensibilité et l'efficacité du pointage de l'antenne mais l'on verra qu'il a aussi d'autres applications.

Avant d'entrer dans le vif du sujet, un chapitre préliminaire présente les éléments nécessaires aux thèmes développés par la suite. Ce chapitre présente plus en détail les caractéristiques du télescope de 30 mètres et les méthodes utilisées pour effectuer et étalonner les observations.

Chapitre 2

Radioastronomie

Ce chapitre essaie avant tout d'établir les éléments nécessaires à la compréhension des deux chapitres suivants, mais il peut aussi être regardé comme une tentative d'introduction à la radioastronomie d'un point de vue pratique. Il est centré sur le radiotélescope de 30 mètres, donc plus particulièrement aux longueurs d'onde millimétrique en antenne unique. Il insiste sur ce qu'il est important de savoir pour pouvoir observer correctement et utiliser le plus efficacement possible le temps d'observation conquis de haute lutte.

Avant de pouvoir utiliser un radiotélescope, il est préférable d'avoir quelques notions sur son architecture et c'est donc par cela que nous allons commencer avant de passer aux procédures d'observation et d'étalonnage proprement dites.

2.1 Architecture d'un radiotélescope

On peut en une première approche diviser un radiotélescope en trois grandes parties qui sont l'antenne, les récepteurs et les spectrographes. Un système informatique contrôle et synchronise ces trois sous-ensembles en fonction des ordres donnés par l'observateur, et assure aussi l'acquisition des données, leur stockage sur support magnétique et l'affichage des résultats.

2.1.1 L'antenne

Les antennes de radioastronomie revêtent des formes variées selon la gamme de longueur d'onde qu'elles sont destinées à étudier. L'emploi d'un réflecteur parabolique est toutefois systématique dès que l'on peut utiliser des techniques optiques ou quasi-optiques pour contrôler et guider les faisceaux avec des miroirs de taille raisonnable. A de rares exceptions près, ces réflecteurs possèdent une ouverture circulaire, ce qui présente l'avantage de donner la même résolution angulaire sur les deux axes, et sont supportés par une monture altazimuthale, la seule utilisable étant donnés leur taille et leur poids.

Le but ici n'est cependant pas de réécrire la théorie des antennes, très complexe quand on fait entrer en ligne de compte les problèmes de tolérance, fondamentaux pour les grands radiotélescopes, et qui requièrent la connaissance de la statistique des défauts en amplitude et en longueur de corrélation. Rappelons simplement que les erreurs de phase font diminuer très rapidement le rendement dès que l'écart type de la surface à sa forme idéale atteint un vingtième à un dixième de la longueur d'onde d'utilisation.

Les structures de grande dimension se déforment sous leur propre poids. Cette déformation dépend de l'orientation, plus précisément de l'élévation pour une monture altazi-

muthale, et son amplitude est très supérieure à la tolérance sur la précision de surface de l'antenne. Pour résoudre ce problème, on fait appel à des structures dites homologues: la forme du réflecteur reste parabolique mais sa distance focale varie ce qui est compensé par un déplacement du récepteur si celui-ci se trouve au foyer primaire, ou du miroir secondaire dans les autres cas.

Pour des raisons techniques, la surface des grands réflecteurs est généralement constituée de nombreux panneaux juxtaposés dont il faut ajuster la position et le 30 mètres n'échappe pas à la règle. Pour tenir l'objectif d'une précision de surface de $70\ \mu$, ce qui explique par ailleurs qu'un système de contrôle thermique combatte les gradients de température, provoqués par exemple par le rayonnement solaire, qui produiraient des déformations très supérieures à la tolérance requise, il faut pouvoir mesurer la forme du miroir primaire avec une erreur de l'ordre de $20\ \mu$. Cette mesure s'effectue par des techniques holographiques qui consistent à réaliser des cartes du diagramme de rayonnement de l'antenne, ce qui exige une très grande dynamique. On verra là qu'il s'agit d'une application un peu à part des détecteurs continuum décrits au chapitre 3, même s'ils n'ont pas été optimisés pour ce genre de travail et peuvent être améliorés.

Signalons au passage que la définition, donnée dans les livres classiques, du rôle d'une antenne comme transition entre un mode de propagation libre et un mode de propagation guidé semble quelque peu dépassée. Dans le domaine millimétrique, on utilise des méthodes quasi-optiques pour diriger le faisceau, et les mêmes termes qu'en astronomie optique définissent la configuration du télescope (Cassegrain, Nasmyth, ...). C'est seulement le cornet du récepteur, lorsqu'il ne s'agit pas d'un mélangeur à structure ouverte, qui effectue la transition vers un mode guidé.

La configuration de la cabine des récepteurs du télescope est de type Nasmyth, ce qui signifie que deux miroirs plans, dont l'un suit l'antenne en élévation, ramènent le faisceau dans un plan horizontal. D'autres composants, des grilles de polarisation et des miroirs séparent le faisceau et le dirigent vers les récepteurs.

Font aussi partie de l'antenne les mécanismes d'entraînement et de contrôle de la position. Pour des antennes à haute résolution, ces mécanismes atteignent un haut degré de sophistication. Dans le cas de l'IRAM, les codeurs angulaires possèdent dix points de résolution par seconde d'arc et la poursuite est assurée par des programmes d'asservissement relativement complexes qui tournent dans des microordinateurs dédiés et qui recalculent les paramètres de commandes des moteurs toutes les 6 ms. Ces programmes doivent par exemple éviter d'exciter les fréquences de résonance de l'antenne mais doivent réagir le plus rapidement possible à la perturbation créée par un changement brusque de la vitesse du vent.

2.1.2 Le ou les récepteurs.

Les récepteurs se divisent essentiellement en deux grandes classes, basées sur des principes de fonctionnement totalement différents: il s'agit d'une part des récepteurs hétérodynes et d'autre part des bolomètres.

Récepteurs hétérodynes

Ces récepteurs sont des mélangeurs dans lesquels un élément non linéaire, diode Schottky ou jonction SIS (supraconducteur-isolant-supraconducteur) génère un battement entre le signal issu d'un oscillateur local, le plus souvent une diode Gunn, à la fréquence ν_{ol} proche de la

fréquence à observer et celui reçu par l'antenne. Ce battement est très fortement amplifié et filtré dans la bande dite de fréquence intermédiaire centrée en ν_{fi} et de largeur totale $\Delta\nu$. Le signal du ciel à la fréquence ν_c produit deux battements de fréquence $\nu_b = |\nu_{ol} \pm \nu_c|$. Le battement supérieur peut être ignoré, ce qui revient à dire que l'on observe le ciel aux fréquences ν_c qui vérifient:

$$\nu_c = \nu_{ol} \pm \nu_b \quad \text{avec} \quad |\nu_b - \nu_{fi}| < \Delta\nu/2 \quad (2.1)$$

La fréquence intermédiaire est comprise entre un et quelques gigahertz sur une bande toujours inférieure à l'octave.

Le radiotélescope dispose en permanence de récepteurs couvrant les bandes de 3 mm ($\approx 80\text{--}116$ GHz), 2 mm ($\approx 130\text{--}183$ GHz), 1.3 mm ($\approx 210\text{--}265$ GHz), et pendant certaines périodes d'hiver la bande de 0.8 mm ($\approx 320\text{--}360$ GHz). La présence d'un récepteur 7 mm ($\approx 40\text{--}45$ GHz), utilisé uniquement pour l'holographie et l'interférométrie à très large base (VLBI), est anecdotique.

Le paramètre essentiel d'un récepteur est son bruit, et tous les efforts visant à le minimiser doivent être entrepris. Ce bruit vient de deux sources principales: la première est le mélangeur proprement dit, la deuxième est le premier amplificateur de fréquence intermédiaire. Cet amplificateur est critique car son bruit est multiplié par les pertes de conversion du mélangeur, qui sont couramment de l'ordre de 6 dB (un facteur 4 en puissance). Ces considérations expliquent que les récepteurs soient installés dans des cryostats: les mélangeurs à diode Schottky et les premiers amplificateurs de fréquence intermédiaire sont maintenus à 15–20 K, quant aux mélangeurs supraconducteurs SIS, ils doivent être refroidis à 3–4 K. Le gain total de la chaîne d'amplification, divisée en une partie froide et une partie à la température ambiante, est énorme: environ 100 dB (un facteur 10^{10} en puissance).

La relation 2.1 montre que le mélangeur déplace deux bandes de fréquence, situées de part et d'autre de l'oscillateur local, vers une gamme plus facilement traitable. En mode spectral, on n'est intéressé sauf exception que par le signal présent dans l'une de ces bandes, appelée bande signal par opposition à la bande image. Dans le cas des récepteurs SIS, il est presque toujours possible d'effectuer un réglage dit simple bande qui déséquilibre la réponse du mélangeur au détriment de la bande supérieure. Cette dissymétrie du comportement est mal comprise mais cela n'empêche pas de la mettre à profit pour améliorer l'efficacité du télescope. Le gain que l'on obtient est particulièrement évident lorsque l'opacité de l'atmosphère, et donc son émission, est beaucoup plus forte dans la bande image que dans la bande signal. C'est par exemple le cas pour la transition $J = 0 \mapsto 1$ de ^{12}CO , située à 115.27 GHz pour laquelle la bande image correspond à une raie forte de l'oxygène. L'autre avantage que présente ce type de réglage est de diminuer les risques de confusion dans les sources très riches; il faut cependant bien se rendre compte que ce risque n'est pas totalement éliminé: le facteur de réjection est le plus souvent compris entre 4 et 20, bien moins que la différence d'intensité entre deux raies dues à des corps distincts et d'abondances différentes.

Les mélangeurs ne sont sensibles qu'à une seule polarisation linéaire, il suffit donc d'une simple grille séparatrice, composant à très large bande, pour diviser le faisceau vers deux récepteurs. Par contre, pour pouvoir observer simultanément avec trois récepteurs, il est nécessaire d'introduire dans le chemin optique un composant sélectif. Sans entrer dans les détails ceci est réalisé au 30 mètres par un système de grilles et de miroirs dont le résultat est une rotation de la polarisation proportionnelle à la fréquence. Il présente le défaut de ne pas permettre d'observer de façon optimale avec toutes les combinaisons de fréquence possible des divers récepteurs, bien que dans la plupart des cas on puisse obtenir un compromis raisonnable.

L'axe de polarisation dans la cabine est au télescope toujours horizontal ou vertical pour un récepteur donné, mais il dépend de l'orientation de l'antenne une fois projeté sur la sphère céleste. Ce phénomène se remarque lorsqu'on observe certains masers pour lesquels l'intensité des raies dépend fortement de l'angle horaire de la source.

Bolomètres

L'autre type important de récepteur est le bolomètre, qui utilise le fait que la résistance d'un cristal de germanium dépend très fortement de sa température lorsqu'il est refroidi à moins de 1 K. On mesure alors directement la variation de température provoquée par le rayonnement absorbé par le cristal, ce qui signifie par ailleurs qu'il ne filtre aucune polarisation.

Le bolomètre est un instrument à très large bande, de l'ordre de 50 GHz pour ceux utilisés au 30 mètres et aucune analyse spectrale plus fine n'est possible puisque l'on transforme directement la puissance reçue en signal continu. Il offre par contre la meilleure sensibilité pour les observations en continuum, de l'ordre de $50 \text{ mJy.s}^{-1/2}$ quand les conditions météorologiques sont favorables.

Récemment, les premiers essais d'un bolomètre multicanal ont été effectués et couronnés de succès. Cet instrument observe simultanément 7 positions situées aux sommets et au centre d'un hexagone; il permet d'effectuer certaines observations en un temps nettement plus court.

Il est aussi possible de construire aussi des récepteurs hétérodynes multicanaux, mais il est peu probable qu'ils soient disponibles au 30 mètres dans un proche avenir.

Autres types de récepteurs

Les progrès des transistors à effet de champ à haute mobilité d'électrons rendent envisageable à moyen terme la disponibilité d'amplificateurs fonctionnant directement à 3 mm. La bande instantanée que l'on pourrait couvrir serait alors de l'ordre de 15 GHz, ce qui devrait se traduire par une amélioration importante de la sensibilité en mode continuum.

2.1.3 Spectrographes

Dans le cas des récepteurs hétérodynes. Le signal, amplifié à des niveaux de puissance où le bruit de l'électronique n'est plus un critère essentiel et ramené dans une bande de fréquence fixe et relativement facile à traiter, peut enfin être analysé plus finement. Les exigences de sensibilité demandent que l'on effectue une analyse spectrale simultanée de toute la bande à observer et non un balayage comme dans un analyseur de spectre de laboratoire. On verra par contre que l'on peut se contenter, à la condition de bien préciser sa définition, d'une dynamique relativement faible.

Les principaux types d'analyseurs utilisés en radioastronomie sont décrits ci-dessous. Cette liste exclut les détecteurs continuum, qui sont un cas à part puisqu'ils ne divisent pas la bande passante en canaux individuels et auxquels un chapitre entier est consacré.

Batteries de filtres

Le principe de ce type d'analyseur de spectre est simple, puisqu'il est équivalent à séparer le signal reçu par autant de filtres électroniques que de canaux souhaités et à mesurer la

puissance à l'aide d'un détecteur à diode. Dans la pratique, le signal est découpé en sous-bandes et ramené par une série de mélanges et de préfiltrages à des fréquences inférieures à 100 MHz. Cette architecture permet de se contenter de filtres simples, avec des largeurs relatives importantes qui ne posent pas de problèmes de stabilité en température et dans le temps.

Il y a l'heure actuelle à l'IRAM deux batteries de 512 filtres de 1 MHz (largeur de la réponse à mi-puissance) et une de 256 filtres de 100 kHz. Ces spectrographes sont assez imposants et gourmands en électricité, mais leur fiabilité est bonne. La plupart des pannes, curieusement plus fréquentes dans la partie logique que dans la partie analogique pourtant plus délicate, sont limitées à des canaux individuels, et n'empêchent pas l'exploitation des spectres obtenus.

Les problèmes posés par les détecteurs à diode, dans le cas particulier des détecteurs continuum, qui mesurent la puissance totale dans la bande du récepteur, sont exposés en détail au chapitre 3. Il s'agit là d'un cas un peu différent car il n'y a que quelques canaux à réaliser, mais les exigences sur les caractéristiques de bruit et de stabilité sont nettement plus sévères.

Corrélateurs

Les corrélateurs sont l'application pratique du théorème de Wiener-Khintchine, qui dit que le spectre de puissance d'un signal est la transformée de Fourier de sa fonction d'autocorrélation. Cela revient à dire que, pour analyser le signal, il faut l'échantillonner aux instants $t_0, t_0 + \tau, \dots, t_0 + n\tau \dots$ où τ est l'inverse de la fréquence d'échantillonnage et calculer le terme de corrélation $c_i, 0 \leq i < m$ correspondant au délai $i\tau$ donné par

$$c_i = \sum_{n=-\infty}^{+\infty} V(t_0 + n\tau)V(t_0 + n\tau + i\tau)$$

où $V(t)$ représente le signal à l'instant t . La transformée de Fourier des c_i donne le spectre du signal dans une bande dont la largeur est limitée à $1/(2\tau)$ par le théorème de Shannon.

La moindre mauvaise valeur dans la fonction d'autocorrélation se traduit par une ondulation sinusoïdale du spectre après la transformation de Fourier, ce qui signifie que toute erreur rend la machine inutilisable.

Le plus étonnant de ce genre d'instrument est que, sous l'hypothèse que le signal à analyser est du bruit gaussien, on peut se contenter d'un échantillonnage extrêmement grossier, sur 1 ou 2 bits. Certes cela introduit une certaine distortion dont la correction, appelée correction de Van-Vleck, est relativement complexe sauf dans le cas du corrélateur à un bit, mais si les seuils de comparaison sont correctement choisis par rapport à la valeur efficace du signal à analyser, la perte de sensibilité n'atteint que 15 % dans le cas d'un corrélateur à 4 niveaux. Les corrélateurs à un bit, qui n'échantillonnent que le signe du signal, souffrent par contre de quelques défauts rédhibitoires, non seulement il font perdre un facteur proche de 2.5 sur le temps d'intégration, mais en plus il est impossible d'étalonner les observations car ils sont théoriquement insensibles au niveau de puissance.

La fréquence d'échantillonnage est limitée en pratique à quelques centaines de mégahertz, l'analyse de bandes larges se fait par concaténation de plusieurs sous-bandes plus étroites, ce qui pose des problèmes si des dérives différentielles de gain entre les sous-bandes apparaissent. Par contre, les corrélateurs sont de loin les analyseurs les plus flexibles: une simple commutation de filtre et de fréquence d'échantillonnage permet de jouer sur la résolution et de s'adapter dans une large mesure à l'objet à observer. On peut aussi obtenir facilement un

très grand nombre de canaux car, une fois le signal échantillonné, la résolution ne dépend que du nombre de délais, c'est à dire de la quantité d'électronique logique, reproductible et sans réglages ni dérives avec le temps ou la température, au contre-exemple qui suit près.

L'expérience à l'IRAM avec un corrélateur d'origine américaine est en effete pour le moins mitigée. Il faut dire que cet appareil exhibe une série impressionnante d'exemples de manque de rigueur de conception. L'électronique de la chaîne de multiplicateurs ignore par exemple totalement les dispersions qu'il peut y avoir sur les composants, ce qui est intolérable sur des circuits qui doivent être reproduits à des milliers d'exemplaires. Le reste de la machine est du même acabit. Depuis peu, un nouveau corrélateur conçu à l'IRAM et construit pour la partie logique autour de circuits de corrélation spécifiquement développés pour la radioastronomie, donne des résultats extrêmement prometteurs bien que toutes les possibilités qu'il offre ne soient pas encore exploitables.

Spectrographes acousto-optique

Dans ce genre d'analyseur, le signal radiofréquence, transformé en onde acoustique dans un cristal appelé cellule de Bragg, diffracte un faisceau laser. L'analyse par une barrette de photodiodes du faisceau dévié donne directement le spectre de puissance du signal électrique.

Ce type de spectrographe possède donc l'avantage d'une très grande simplicité théorique et d'un nombre de composants très réduit, une seule cellule de Bragg pouvant analyser une bande de 1 GHz avec une résolution de l'ordre de 1 MHz. Cependant la réalisation demande beaucoup de soins pour obtenir la stabilité requise par les applications radioastronomiques. Le fait qu'il y ait peu de composants est un argument en faveur de la fiabilité, mais il faut tenir compte du fait que certains, comme la source laser, sont relativement fragiles et que les pannes affectent toute la machine.

Le chapitre 4 donne plus de détails sur les caractéristiques, avantages et inconvénients du spectrographe acousto-optique réalisé en collaboration avec l'Observatoire de Meudon et installé à l'IRAM.

2.2 Méthodes d'observation

Le signal analysé par les spectrographes est la somme de quelques termes qui sont: le bruit du récepteur, les pertes dans la cabine des récepteurs et dans les lobes arrières de l'antenne, l'émission de l'atmosphère et enfin le signal astronomique, sauf exception de loin le plus faible.

Pour donner un ordre de grandeur, disons qu'on observe couramment des raies qui représentent à leur maximum moins de 10^{-4} de la densité spectrale de puissance à l'entrée des analyseurs. Dans le cas des galaxies à fort décalage vers le rouge, on a même poussé l'intégration jusqu'à des valeurs inférieures à 10^{-5} . Dans ces conditions, la seule méthode possible consiste à utiliser des techniques de modulation du signal astronomique et de détection synchrone de cette modulation.

Les méthodes de modulation du signal à observer se résument en deux branches: modulation par commutation de lobe et modulation par commutation de fréquence. Elles sont décrites ci-dessous avant de passer à une analyse plus détaillée des caractéristiques de la détection synchrone.

2.2.1 Modulation par commutation de lobe

Cette méthode consiste à déplacer le lobe de l'antenne entre la source à observer et un point proche sur le ciel, appelé position de référence et libre de toute émission ou supposé tel, quelquefois difficile à dénicher dans la cas de sources très étendues.

Trois mécanismes distincts permettent de réaliser ce type d'observations à l'IRAM:

1. La méthode qui donne les meilleurs résultats consiste à commuter la position du miroir secondaire entre deux positions symétriques par rapport à l'axe optique de l'antenne. Ce mécanisme, appelé « wobbler » en anglais, limite la distance entre la source et la position de référence à 4 minutes d'arc, la période de commutation peut par contre être relativement courte. Pour les observations de raies spectrales, la perte de temps d'intégration due au mouvement du secondaire est de comprise entre 5 et 10 %. Pour le moment, l'axe qui sépare les lobes correspond à un mouvement en azimuth, ce qui signifie que la position de référence bouge sur le ciel au cours des observations. L'expérience a de plus montré qu'il est nécessaire d'alterner les cycles dans lesquels la source se retrouve dans le lobe gauche avec ceux dans lesquels elle se retrouve dans le lobe droit ce qui signifie que l'on utilise deux positions de référence symétriques par rapport à la source. Cette alternance, qui signifie un mouvement de l'antenne, peut être très lente et n'affecte pas de façon sensible l'efficacité du télescope.
2. Une autre possibilité, utilisée exclusivement en mode continuum, est fournie par le « beam-switch », qui est une roue avec deux pales placée devant le miroir quaternaire. Lorsque le faisceau, vu depuis la cabine, rencontre une des pales, il se retrouve décentré sur le miroir secondaire. L'écart angulaire entre les lobes vaut environ 90 secondes d'arc et la direction projetée sur le ciel de l'axe les séparant dépend de l'orientation de l'antenne. La fréquence de commutation est élevée, de l'ordre de 7 Hz, mais la proportion de temps perdue entre les phases d'intégration est forte: elle dépasse un peu 50 %.
3. La procédure dite de « position-switch », qui ne présente d'intérêt qu'en mode spectral, est la plus simple. Elle consiste tout simplement à déplacer l'antenne entre la source et la position de référence. Il est évident qu'on ne peut pas utiliser des cycles très rapides, le moindre mouvement de l'antenne demandant quelques secondes: les cycles durent d'habitude une minute, ce qui est très exigeant sur la stabilité du système et de l'atmosphère. C'est la seule méthode applicable pour les sources trop étendues pour utiliser le « wobbler », la position de référence peut être choisie fixe sur le ciel et être très éloignée, mais cela a pour conséquence de dégrader sérieusement la qualité des spectres obtenus, surtout si la différence d'élévation entre les deux positions est importante.

Aucune de ces méthodes ne permet en théorie de cartographier en mode continuum une source étendue. Il existe pour combler cette lacune une technique de reconstruction d'images [25]. Elle est surtout utilisée avec les bolomètres.

2.2.2 Modulation par commutation de fréquence

Cette méthode, qui n'est utilisable que pour les récepteurs hétérodynes en mode spectral consiste à faire varier périodiquement la fréquence de l'oscillateur local. Dans ce cas la raie à observer est présente en permanence dans le signal, simplement la même fréquence ciel

se déplace dans le spectre observé, ce qui par rapport à la commutation de lobe présente l'avantage de réduire le temps d'intégration nécessaire pour obtenir un niveau de bruit donné par un facteur deux, et les problèmes de recherche d'une position de référence dans les sources étendues sont éliminés.

Cette technique ne permet cependant d'observer que des raies étroites: la limite théorique est qu'on ne peut dépasser la moitié de la bande passante du récepteur, la limite pratique dépend de l'écart de fréquence que sont prêts à supporter le récepteur et son oscillateur local, tout au plus quelques dizaines de mégahertz.

Autant l'avouer, cette méthode ne donne pas toujours de résultats satisfaisants au 30 mètres. Des améliorations récentes la rendent exploitable lorsqu'on se contente d'utiliser un seul récepteur.

2.2.3 Détection synchrone

Il a donc été dit que l'on effectue deux mesures successives et que l'on évalue la différence entre ces deux mesures. Idéalisons ceci en supposant que la première valeur est le résultat de l'intégration entre les instants $t = -T$ et $t = 0$, et la deuxième entre $t = 0$ et $t = T$. Supposons que l'on affecte le poids -1 à la première et 1 à la seconde. Au point de vue des fluctuations de puissance, ce procédé est équivalent à réaliser un filtrage dont la réponse en fréquence est la transformée de Fourier de la fonction $f(t)$ qui vaut -1 entre $-T$ et 0 , 1 entre 0 et T , et zéro partout ailleurs. Cette transformée de Fourier $f(\omega)$ a pour carré du module:

$$|f(\omega)|^2 = \frac{16 \sin^4(\omega T/2)}{\omega^2 T^2} \quad (2.2)$$

qui est équivalent à $\omega^2 T^2$ dans la limite où ω tend vers 0 .

Lorsqu'on répète le cycle (par exemple $\{-1, 1, -1, 1\}$), on améliore la sélectivité du filtre équivalent. Lorsqu'on rend le filtre symétrique ($\{-1, 1, 1, -1\}$), on augmente la puissance de ω dans la limite des basses fréquences et il est clair que les dérivées linéaires en fonction du temps sont éliminées. Dans le cas d'un très grand nombre de cycles, la différence entre les deux types de cycle diminue considérablement et la réponse devient celle d'un filtre extrêmement sélectif ne laissant passer que les fréquences telles que ωT soit un multiple impair de π . L'analogie avec les filtres numériques, du moins ceux de type non récursif, qui consistent à effectuer la moyenne pondérée d'une série d'échantillons à chaque pas de temps, est très forte.

On voit que ces filtres possèdent toute une série de zéros correspondant aux cas où ωT est un multiple de 2π , et on peut donc rejeter efficacement une perturbation de pulsation connue. L'expérience montre qu'il est fondamental de considérer ce problème en mode continuum, le plus mauvais choix correspondant $\omega T = \pi$, qui réalise une détection synchrone de la perturbation, quelquefois plus forte que la source à observer.

2.2.4 Etalonnage

Bien que le mot « calibration » soit entré dans le vocabulaire de la communauté des radioastronomes et qu'il existe en français dans un sens bien différent, il s'agit d'un anglicisme et il semble préférable d'utiliser la traduction officielle. La principale réserve que l'on pourrait émettre est que le terme étalonnage peut évoquer un degré de précision que l'on est loin d'atteindre.

Le but de l'étalonnage est de rendre le résultat des observations indépendant des conditions dans lesquelles elles ont été réalisées [26]. Il compense les différences de gain entre les canaux des spectrographes, la forme de la bande passante du récepteur et l'absorption atmosphérique. Entre des antennes différentes, il faut par contre aussi tenir compte de la différence de coefficient de couplage avec la source. L'objectif est d'obtenir une précision de l'ordre de 10 %.

En radioastronomie, on a coutume de déterminer les flux reçus en température de brillance équivalente d'un corps noir dans l'approximation de Rayleigh-Jeans. Cette tradition, héritée du domaine des radars et communications radio, n'est pas toujours valable aux longueurs d'onde millimétriques, surtout dans le cas de nuages froids. Il se trouve aussi que ce système d'unité est particulièrement commode vu la méthode d'étalonnage utilisée: celle du « chopper wheel » [27, 28], qui consiste à placer successivement devant le récepteur deux corps noirs à deux températures distinctes, ou charges pour employer le terme consacré. La première de ces charges est à la température ambiante T_{amb} (290–300 K), et la deuxième est refroidie par de l'azote liquide (77 K) et sa température effective, parfois plus élevée à cause de problèmes de couplage du récepteur à la charge et de pertes dans le chemin optique à température ambiante, sera notée T_{fr} . Les puissances mesurées dans les deux cas, notées P_{amb} et P_{fr} permettent de calculer la température de bruit du récepteur T_{rec} par la relation:

$$T_{\text{rec}} = \frac{T_{\text{amb}}P_{\text{fr}} - T_{\text{fr}}P_{\text{amb}}}{P_{\text{amb}} - P_{\text{fr}}}$$

On effectue aussi une troisième mesure sur le ciel hors source et la puissance mesurée P_{hs} donne une température équivalente T_{hs} qui vaut:

$$T_{\text{hs}} = \frac{T_{\text{fr}}(P_{\text{amb}} - P_{\text{hs}}) + T_{\text{amb}}(P_{\text{hs}} - P_{\text{fr}})}{P_{\text{amb}} - P_{\text{fr}}}$$

Cette température équivalente est la somme de deux contributions selon la formule:

$$T_{\text{hs}} = \eta_c T_b + (1 - \eta_c) T_p$$

où T_b est la température de brillance de l'atmosphère, T_p est la température effective des pertes, prise comme une moyenne pondérée de la température de la cabine des récepteurs et de la température extérieure, pour tenir compte de leur répartition entre l'optique et les lobes arrières de l'antenne. Le coefficient η_c , qui mesure le couplage avec le ciel (« forward efficiency »), dépend de l'antenne et du récepteur utilisé. En toute rigueur, il varie aussi avec l'élévation mais assez faiblement pour pouvoir le considérer comme constant.

On a vu que les récepteurs hétérodynes reçoivent en fait deux bandes centrées aux fréquences signal ν_s et image ν_i avec des rendements différents. Si l'on note G_i le rapport de gain entre les bandes image et signal dans lesquelles les températures de brillance valent respectivement T_{b_i} et T_{b_s} , on peut écrire:

$$T_b = \frac{T_{b_s} + G_i T_{b_i}}{1 + G_i}$$

un modèle, dont le seul paramètre variable est quantité de vapeur d'eau précipitable, détermine quatre quantités qui sont les opacités au zénith τ_s et τ_i et les températures effectives de l'atmosphère T_{a_s} et T_{a_i} et qui vérifient

$$T_{b_s} = T_{a_s}(1 - \exp(-\tau_s A_m)) \quad \text{et} \quad T_{b_i} = T_{a_i}(1 - \exp(-\tau_i A_m))$$

Dans cette expression, A_m représente le nombre de masses d'air sur la ligne de visée, égal à de petites corrections près à l'inverse du sinus de l'élévation.

Profitons de l'occasion pour indiquer ici que l'hypothèse implicite d'uniformité de l'opacité de l'atmosphère sur chaque bande est un peu hardie lorsqu'on l'on observe en large bande à une fréquence correspondant à l'aile d'une raie atmosphérique; c'est le cas pour la première transition de CO à 115.27 GHz alors qu'une raie intense de l'oxygène est située environ 2.7 GHz plus haut. L'étalonnage au centre de la bande, le plus important, n'est cependant pas affecté par cette approximation.

On définit la température étalon T_e comme la température de la source astronomique qui donnerait la même puissance que la charge chaude:

$$T_e = \frac{(T_{\text{amb}} - T_{\text{hs}})(1 + G_i)}{\eta_s \exp(-\tau_s A_m)}$$

où η_s représente le couplage avec la source. Cette dernière valeur est soit égale à η_c , et en ce cas on dit que l'étalonnage est effectué en température d'antenne corrigée de l'absorption atmosphérique et des pertes ohmiques (notée T_R^* dans la littérature depuis Kutner et Ulich [28]), soit lui est inférieure et mesure le couplage du télescope avec une source d'extension angulaire donnée. Dans le cas où l'on utilise le couplage du lobe principal « main-beam efficiency », il faut tenir compte du fait que, à haute fréquence, il varie sensiblement avec l'élévation. Mais comme η_s n'est qu'un facteur d'échelle, il peut être modifié à posteriori par une simple multiplication du spectre.

Enfin le paramètre le plus important pour l'observateur est la température système T_{sys} donnée par:

$$T_{\text{sys}} = \frac{T_e P_{\text{hs}}}{P_{\text{amb}} - P_{\text{hs}}}$$

elle mesure la qualité de l'ensemble télescope, récepteur et atmosphère: le temps d'intégration, ou temps de pose dans le langage des opticiens, pour obtenir une sensibilité donnée est proportionnel à T_{sys}^2 .

2.2.5 Linéarité

Toute erreur dans les mesures de puissance lors de l'étalonnage se traduira par une erreur de flux dans les résultats des observations. Il est donc important de s'assurer que les erreurs de linéarité du système restent dans des limites tolérables. En reprenant les formules précédentes, on peut réécrire:

$$T_e = (P_{\text{amb}} - P_{\text{hs}}) \left(\frac{T_{\text{amb}} - T_{\text{fr}}}{P_{\text{amb}} - P_{\text{fr}}} \right) \left(\frac{1 + G_i}{\eta_s \exp(-\tau_s A_m)} \right)$$

en introduisant maintenant P_{ss} comme la puissance reçue sur la source, la température T_s de la source astronomique sera donnée par:

$$T_s = (P_{\text{ss}} - P_{\text{hs}}) \left(\frac{T_{\text{amb}} - T_{\text{fr}}}{P_{\text{amb}} - P_{\text{fr}}} \right) \left(\frac{1 + G_i}{\eta_s \exp(-\tau_s A_m)} \right)$$

en se ramenant directement aux mesures de puissances. Signalons qu'ici P_{hs} n'est pas le résultat de la même mesure que lors du calcul de T_e , mais les valeurs sont très proches sauf en cas d'instabilité de l'atmosphère ou de l'électronique qui compromet irrémédiablement la qualité de l'étalonnage.

Compte tenu des niveaux de puissance auxquels on travaille, la linéarité de la chaîne d'amplification ne pose pas de problèmes majeurs jusque dans la distribution des signaux de fréquence intermédiaire: il faut tout de même y apporter un peu d'attention étant donné le nombre d'étapes successives et le niveau de précision que l'on prétend atteindre. Pour donner une idée, disons que l'on a déterminé expérimentalement que, pour les amplificateurs monolithiques bon marché apparus ces dernières années comme les séries MSA d'Avantek, il faut travailler au plus 20 dB en dessous du point de saturation spécifié par le fabricant.

Il en va tout autrement lorsqu'on a besoin d'un amplificateur de puissance pour délivrer le signal à la cellule de Bragg d'un spectrographe acousto-optique et surtout lorsqu'on utilise des détecteurs à diode. En effet, dans ce dernier cas, ce que l'astronome appelle un détecteur linéaire est un circuit hautement non linéaire, en fait quadratique, pour l'électronicien.

On peut modéliser ceci de la façon suivante: l'instrument de mesure de puissance fournit une valeur numérique $N(P)$ fonction de la puissance qui est en première approximation proportionnelle à P avec un terme d'erreur petit et variant lentement avec P :

$$N(P) = \alpha P(1 + \epsilon(P))$$

en remplaçant les puissance par leur valeur numérisée dans l'expression de T_s , on obtient, après simplification:

$$T_s = \frac{(P_{ss} - P_{hs} + P_{ss}\epsilon(P_{ss}) - P_{hs}\epsilon(P_{hs}))(T_{amb} - T_{fr})(1 + G_i)}{(P_{amb} - P_{fr} + P_{amb}\epsilon(P_{amb}) - P_{fr}\epsilon(P_{fr}))\eta_s \exp(-\tau_s A_m)}$$

dans l'immense majorité des cas P_{ss} et P_{hs} sont tellement proches que l'on peut supposer $\epsilon(P_{ss}) = \epsilon(P_{hs})$, par contre, la différence entre P_{amb} et P_{fr} est forte, puisque leur rapport est souvent compris entre 1.5 et 2.5, une valeur proche de 2 étant la plus courante. Réécrivons donc T_s en tenant compte de ces remarques et du fait que l'arbitraire dans le choix de la fonction ϵ nous autorise à poser $\epsilon(P_{hs}) = 0$:

$$T_s = (P_{ss} - P_{hs}) \left(\frac{T_{amb} - T_{fr}}{P_{amb} - P_{fr} + P_{amb}\epsilon(P_{amb}) - P_{fr}\epsilon(P_{fr})} \right) \left(\frac{1 + G_i}{\eta_s \exp(-\tau_s A_m)} \right)$$

l'erreur relative $\epsilon(T_s)/T_s$ du résultat de l'observation devient:

$$\frac{\epsilon(T_s)}{T_s} \approx \frac{P_{fr}\epsilon(P_{fr}) - P_{amb}\epsilon(P_{amb})}{(P_{amb} - P_{fr})}$$

dans tous les cas, P_{hs} est plus proche de P_{fr} que ne l'est P_{amb} , le contraire signifierait $T_b > T_{amb}$, donc une atmosphère très opaque et très chaude, en pratique inutilisable pour les observations. Dans le cas de la plupart des circuits électroniques, où l'erreur de linéarité croît très vite dès que l'on s'approche de la saturation, cette expression serait dominée par la contribution du terme en $\epsilon(P_{amb})$; dans le cas d'un détecteur à diode cette caractéristique n'est pas du tout évidente, mais on peut la justifier par le fait qu'elle doit avoir une variation lente sur la gamme d'utilisation et qu'il est fréquent que P_{hs} soit égal à P_{fr} à 25 % près alors que P_{amb}/P_{fr} est souvent compris entre 1.5 et 2.5. On obtient alors:

$$\frac{\epsilon(T_s)}{T_s} \simeq \frac{-\epsilon(P_{amb})}{(1 - P_{fr}/P_{amb})}$$

dans le cas courant où le facteur $P_{amb}/P_{fr} \simeq 2$, l'erreur sur l'étalonnage vaut le double de l'erreur de mesure de l'instrument. Toute erreur de linéarité supérieure à 3 % sera une

contribution importante au budget d'erreur, il faut donc essayer de rester systématiquement en dessous de cette limite. On en conclut aussi que l'étalonnage, contrairement à ce que pourrait faire penser le terme, amplifie les erreurs des instruments.

Nous avons jusqu'ici ignoré l'effet de l'erreur de linéarité dans l'expression de T_e alors qu'il est évident qu'elle intervient aussi par l'intermédiaire du calcul de T_{hs} et de l'opacité. La raison en est de nouveau que P_{hs} est suffisamment proche de P_{fr} pour que l'erreur introduite soit faible. Il est possible d'obtenir une estimation correcte de T_e et des paramètres atmosphériques alors que l'étalonnage est faux!

2.2.6 Dynamique

La plus grande variation de la puissance que l'on peut attendre à la sortie du récepteur correspond à la différence entre la charge chaude d'étalonnage et le ciel aux fréquences où l'atmosphère absorbe peu dans de bonnes conditions météorologiques. Ce rapport Y vaut:

$$Y = \frac{T_{rec} + T_{amb}}{T_{rec} + \eta_c T_a + (1 - \eta_c) T_p}$$

presque tous les termes peuvent être pris comme constants en première approximation: $T_{amb} = 300$ K, $T_p \simeq 280$ K et $\eta_c = 0.9$. La seule variable restante est donc la température de bruit du récepteur; supposons que l'on a un récepteur excellent avec $T_{rec} = 40$ K, on obtient alors $Y \simeq 5$.

Ces chiffres pourraient donner l'impression que la dynamique requise est ridicule. Mais il faut faire attention à sa définition, c'est la dynamique sur laquelle on ne veut pas voir d'augmentation de l'incertitude de mesure due au bruit propre des détecteurs ou des spectrographes. Or, l'incertitude relative de mesure pour un temps d'intégration τ vaut $1/\sqrt{B_{ac}\tau}$, où B_{ac} est la largeur d'autocorrélation du filtre utilisé, que nous appellerons aussi bande équivalente de fluctuation ou bande de fluctuation tout court. Si on note $R(\nu)$ la réponse en fréquence d'un filtre, qu'il soit réalisé physiquement où que l'on utilise son équivalent dans le cas des corrélateurs ou des spectrographes acousto-optiques, B_{ac} vaut:

$$B_{ac} = \frac{(\int R(\nu) d\nu)^2}{\int R^2(\nu) d\nu}$$

par exemple dans le cas d'un filtre de profil gaussien, cette largeur vaut $\sqrt{\frac{\pi}{2 \text{Log } 2}} \simeq 1.5$ fois la largeur à mi-hauteur. Signalons à ce sujet que beaucoup d'astronomes utilisent la largeur à mi hauteur des canaux du spectrographe et non sa bande de fluctuation lorsqu'ils calculent le bruit théorique. C'est théoriquement faux, mais parfaitement valide quand on sait que dans bon nombre de cas, on est obligé de lisser les spectres par sommation de canaux voisins, ce qui a pour effet de donner des filtres équivalents quasiment carrés.

Pour un filtre de largeur 1 MHz à mi-hauteur, la bande de fluctuation sera proche de 1.5 MHz, et l'incertitude de mesure vaut alors 1/1250 de la valeur mesurée en une seconde. Comme tous les canaux ne travaillent pas au même niveau, une ondulation dans la bande du récepteur d'un facteur 2 pouvant être considérée comme faible, une dynamique d'acquisition de l'ordre de 15000 en une seconde peut être considérée comme le strict minimum.

Elle atteint 100000 pour une pose de trente secondes, durée couramment utilisée en mode « position-switch », mais dans ce cas les problèmes de stabilité, abordés au paragraphe suivant, sont les plus sévères. Il vaut mieux bien sûr que cette dynamique soit plus élevée pour minimiser les problèmes stupides lors de l'utilisation quotidienne du télescope.

Pour les détecteurs continuum, mesurant la puissance sur toute la bande de de 500 MHz ou 1 GHz, mais ne souffrant évidemment pas des problèmes d'ondulation et n'utilisant jamais des temps d'intégration supérieurs à la seconde, il faut multiplier ce chiffre par 10.

2.2.7 Stabilité

Des ordres de grandeur du signal astronomique attendu et des méthodes nécessaire pour arriver à l'extraire, on déduit que la stabilité de tous les maillons de la chaîne d'acquisition du télescope est un paramètre fondamental. Il faut tout de suite faire la distinction entre les observations en mode spectral et en mode continuum.

En effet dans le premier cas, la raie apparaît par contraste entre des canaux différents mais acquis simultanément par les spectrographes. Cela signifie que de petites variations globales d'amplitude du signal sont rejetées par la procédure; les dérives différentielles sur la bande observée dégradent par contre sérieusement la qualité des spectres. On appelle ligne de base le résidu de ces dérives et l'une des premières étapes de la réduction des données consiste à l'éliminer, le plus souvent en l'assimilant à un polynôme de degré le plus faible possible, mais aussi, lorsqu'elle possède une structure périodique évidente, par assimilation à une sinusoïde ou par interpolation des composantes de Fourier aberrantes. Les défauts de ligne de base que l'on peut accepter dépendent tellement de l'objet à observer qu'on ne peut donner qu'une règle générale: il faut que sa structure ne présente que des composantes nettement plus larges que la raie à observer.

En mode continuum, au contraire, il est impossible de distinguer les variations de puissance dues à la source astronomique de celles dues aux instabilités de l'atmosphère, des récepteurs ou des détecteurs. Il en résulte que la période de modulation doit être beaucoup plus courte qu'en mode spectral et d'autres problèmes peuvent surgir, par exemple une modulation de la puissance au rythme de la pompe du circuit d'hélium du récepteur. Un choix judicieux de la durée d'intégration permet, comme indiqué au paragraphe 2.2.3, de minimiser ces problèmes.

Dans tous les cas, diminuer la fréquence de commutation se traduira tôt ou tard par une dégradation des caractéristiques du système; l'incrémenter provoquera une augmentation de la proportion de temps mort consacrée au mouvement de l'antenne, du secondaire, etc. Il existe donc une durée de cycle optimale qu'il est important de connaître et qui peut être estimée par la méthode qui suit.

2.2.8 Variances d'Allan

La variance d'Allan [29] fut introduite dans les années 1960 pour mesurer la stabilité des horloges atomiques. Sous une forme légèrement modifiée, c'est l'outil qui permet de déterminer les limitations du système en matière de stabilité.

Dans la pratique, on évalue la variance d'Allan en acquérant un grand nombre N d'intégrations courtes de même durée τ . Notons $x_i, i = 1, \dots, N$ les valeurs acquises. On les somme ensuite par groupe de p points adjacents, ce qui donne une nouvelle série d'échantillons $e_{pj} = \sum_{i=pj-p+1}^{pj} x_i$, avec $j = 1, \dots, E(N/p)$ où $E(N/p)$ est la partie entière de N/p . Comme dernière étape on évalue la moyenne du carré de la différence d'échantillons successifs, c'est la variance d'Allan pour un cycle de durée $2p\tau$:

$$\sigma_A(2p\tau) = \overline{(e_{p(j+1)} - e_{pj})^2}$$

Une fois ce calcul effectué, on trace sur un diagramme avec échelles logarithmique $\sigma_A(t)$ en fonction de t . Pour un système parfait analysant du bruit gaussien, la pente doit être égale à -1 . Au delà d'une certaine durée, on observe à cause de phénomènes physiques inévitables causant des instabilités et des dérives que la variance d'Allan s'éloigne de cette courbe idéale. Le rapport entre la variance d'Allan pour une durée de cycle donnée et sa valeur idéale déduite du comportement dans la limite des faibles durées représente le facteur de perte de temps d'intégration dû à ces imperfections.

A partir des expressions précédentes, on peut élaborer un certain nombre de variantes: par exemple dans le cas d'un spectrographe, on peut éliminer les fluctuations globales de gain en normalisant par la puissance totale reçue. Ces fluctuations ne se traduiraient que par une ligne de base constante et ne sont pas gênantes en pratique, elles sont en fait inévitables en raison du mouvement de la source en élévation et de la variation de la contribution de l'atmosphère qui en résulte.

Des exemples de variance de Allan sont présentés aux chapitres 3 et 4. Ces graphiques ont systématiquement été normalisés de façon à montrer directement la bande équivalente de fluctuation B_{ac} , c'est à dire que $\sigma_A(t) = 1/B_{ac}t$.

Chapitre 3

Continuum

3.1 Motivation

Pour un certain nombre de procédures d'observation, on cherche à mesurer la puissance totale reçue dans la bande passante du récepteur.

C'est le cas lorsqu'on cherche à déterminer l'intensité de rayonnement d'une source émettant un rayonnement continu, thermique par exemple.

C'est aussi le cas d'un certain nombre de procédures permettant d'évaluer les paramètres de l'antenne et de l'atmosphère. Ces procédures sont auxiliaires dans la mesure où elles ne donnent pas de résultats scientifiques en elles-mêmes mais constituent des intermédiaires indispensables à l'interprétation des observations effectuées avec le télescope.

C'est enfin le cas de mesures purement techniques sur les récepteurs et la chaîne d'acquisition du télescope.

Les observations et mesures sont le plus souvent effectuées en commutant rapidement le lobe de l'antenne entre une position sur la source et une position de référence située à plusieurs lobes de la source. Cette méthode, comme expliqué dans le chapitre sur les techniques d'observation, permet d'éliminer le bruit basse fréquence dû à plusieurs sources: fluctuation de la transmission et de l'émission atmosphériques, variation de gain des récepteurs et bruit rose des détecteurs proprement dit.

Les plus importantes de ces procédures sont décrites ci-dessous.

3.1.1 Observations en continuum

Il s'agit essentiellement de mesurer le flux moyen émis dans la bande du récepteur par une source, par exemple pour étudier sa variabilité dans le temps [24]. Deux points sont primordiaux dans ce cas-là: que l'étalonnage soit correct et le bruit aussi proche que possible de la limite théorique.

3.1.2 Pointage

La procédure de pointage consiste à effectuer un balayage en azimuth et en élévation autour de la position estimée d'une source intense et proche de l'objet à observer. On en déduit ensuite les erreurs de positionnement de l'antenne que l'on introduit comme corrections dans le logiciel de poursuite. La faible taille du lobe de l'antenne de 30 m, environ 10 secondes d'arc à 230 GHz, rend le pointage particulièrement critique.

La source de pointage doit être ponctuelle ou tout au plus d'une taille comparable au lobe de l'antenne, et suffisamment intense pour que son émission soit détectable en moins d'une seconde d'intégration.

La plupart des sources utilisées pour le pointage possèdent une émission continuum, par exemple les quasars, on peut donc les observer quelle que soit la fréquence à laquelle est réglé le récepteur. Les exceptions à cette règle sont les sources peu étendues à très forte émission comme peuvent l'être les masers (essentiellement **SiO** à 86.24 GHz) et certaines enveloppes stellaires, utilisables lorsqu'une transition assez forte tombe dans la bande d'un des récepteurs. Dans bien des régions du ciel, il faut se contenter d'objets faibles, à la limite des possibilités de détection. Dans ce cas là, il est fondamental que le bruit et la stabilité du détecteur ne soient pas le facteur limitant la sensibilité du télescope.

3.1.3 Focalisation

Cette procédure permet de vérifier que le faisceau est correctement focalisé sur les récepteurs et, le cas échéant, d'ajuster la position du miroir secondaire pour maximiser le rendement de l'antenne et sa résolution.

L'erreur de focalisation semble être liée à des phénomènes thermiques dans l'antenne, sa variation est lente et la procédure n'a pas besoin d'être répétée aussi souvent que le pointage. De plus la source utilisée pour effectuer cette mesure n'a pas besoin d'être proche de l'objet à observer. C'est heureux car les sources suffisamment puissantes pour permettre une mesure précise de l'erreur de positionnement du miroir secondaire se réduisent pour l'essentiel aux planètes et à quelques quasars.

Dans ce cas comme dans celui du pointage, on demande au détecteur une excellente stabilité et un bruit aussi faible que possible.

3.1.4 Efficacité de l'antenne et opacité de l'atmosphère

Ces deux paramètres sont évalués simultanément comme le résultat de la mesure de la puissance P reçue par le détecteur en fonction de l'élévation (skydip) qui est théoriquement de la forme:

$$P \propto T_r + (1 - \eta_c)T_p + \eta_c T_a \times (1 - e^{-\tau_z A_m})$$

dans cette formule simplifiée (elle est valable seulement pour un récepteur simple bande ou lorsque l'opacité de l'atmosphère est la même dans les deux bandes), T_r représente la température de bruit du récepteur, η_a le coefficient de couplage de l'antenne avec le ciel, T_p la température effective des pertes ohmiques dans la cabine des récepteurs et dans les lobes arrière de l'antenne, T_a la température effective de l'atmosphère, τ_z l'opacité de l'atmosphère au zénith et A_m le nombre de masses d'air sur la ligne de visée. Ce dernier terme est essentiellement égal à l'inverse du sinus de l'élévation, à de petites corrections près qui prennent en compte la sphéricité de l'atmosphère [30, page 166]. Les paramètres T_r et η_c sont fonction du récepteur utilisé et de la fréquence à laquelle il est réglé; T_p , T_a et τ_z dépendent en plus des conditions météorologiques.

La puissance sortant du récepteur représente une mesure de l'émissivité du ciel aux différentes élévations, laquelle est interprétée comme température par interpolation (ou extrapolation dans certains cas) de la puissance mesurée sur les charges chaude et froide d'étalonnage. Il est clair que dans ce cas, les erreurs de linéarité de la mesure, qui sont principalement celles des détecteurs, sont critiques et doivent être minimisées.

3.1.5 Coefficient de couplage du lobe principal

Le lobe principal de l'antenne a une forme proche d'une gaussienne de diamètre angulaire à mi-puissance θ_{mp} . Le coefficient de couplage $\eta(\theta)$ avec une source de brillance uniforme et de diamètre angulaire θ vaut:

$$\eta(\theta) = 2\eta(\theta_{mp}) \left[1 - e^{-\frac{\theta^2}{\theta_{mp}^2}} \right]$$

On appelle coefficient de couplage du lobe principal η_p la quantité $2\eta(\theta_{mp})$. Sa détermination est cruciale dans l'obtention d'un étalonnage correct, permettant une comparaison des résultats obtenus par différents radiotélescopes. La différence entre η_c et η_p représente le couplage avec les lobes secondaires, inévitables puisqu'une antenne de taille finie ne peut avoir un lobe rigoureusement gaussien, sans parler des problèmes de qualité de la surface réfléchissante, de la précision du positionnement des panneaux qui la constituent, du blocage par le miroir secondaire et par les pieds qui le supportent.

La valeur de η_p est calculée le plus souvent par la mesure soigneusement étalonnée d'une planète, source intense de taille parfaitement connue. L'expérience montre qu'il faut choisir une source soit bien plus grande que θ_{mp} , soit nettement plus petite pour obtenir des résultats cohérents. Les mesures sur une source de taille proche de θ_{mp} sont en effet très sensibles à la scintillation et aux erreurs de pointage et de poursuite.

Mis à part η_c et η_p , on utilise aussi le coefficient de couplage avec la Lune η_l . Ce dernier, qui tient compte des premiers lobes secondaires, est surtout utilisé pour les observations de sources très étendues comme les galaxies proches.

3.1.6 Mesures techniques

Il s'agit essentiellement de mesure de stabilité des récepteurs et des divers composants de la chaîne d'acquisition du télescope. Bien que pour la plupart maillons de cette chaîne, on utilise des techniques éprouvées qui ne posent généralement aucun problème, les stabilités de gain nécessaires sont telles qu'aucun instrument commercial ne fournit des mesures suffisamment précises.

3.2 Caractéristiques requises

L'objectif est de mesurer la puissance sortant du récepteur avec suffisamment de précision pour atteindre la limite théorique de sensibilité donnée par le bruit du récepteur. Pour une bande passante B de 1 GHz et un temps d'intégration τ de une seconde, l'incertitude relative vaut $1/\sqrt{B\tau} \approx 1/30000$. Il faut donc que le niveau de signal sortant du détecteur soit supérieur à 100000 fois son bruit propre dans un hertz de bande. La gamme de fréquences dans laquelle le détecteur doit travailler dans le cas de l'IRAM est 80 à 1100 MHz.

Le signal issu d'un récepteur de radio-télescope peut être assimilé à du bruit gaussien, dont les excursions en tension atteignent fréquemment jusqu'à ± 3 fois la valeur quadratique moyenne, qui sera notée V_{eff} . Calculons, pour fixer les idées, la probabilité $P_V(|v| > 3V_{\text{eff}})$ d'obtenir une valeur supérieure à $3V_{\text{eff}}$ et la fraction $P_W(|v| > 3V_{\text{eff}})$ de puissance correspondante:

$$P_V(|v| > 3V_{\text{eff}}) = \frac{2}{\sqrt{2\pi} V_{\text{eff}}} \int_{3V_{\text{eff}}}^{+\infty} \exp\left(\frac{-V^2}{2V_{\text{eff}}^2}\right) dV \approx 0.27\% \quad (3.1)$$

$$P_W(|v| > 3V_{\text{eff}}) = \frac{2}{\sqrt{2\pi} V_{\text{eff}}^3} \int_{3V_{\text{eff}}}^{+\infty} V^2 \exp\left(\frac{-V^2}{2V_{\text{eff}}^2}\right) dV \approx 2.9\% \quad (3.2)$$

Cette estimation, bien qu'idéalisée, est une bonne approximation de la réalité. Il indique que si un composant sature pour une tension inférieure à $3V_{\text{eff}}$, l'étalonnage sera faussé et l'objectif de 10% de précision difficilement réalisable.

Les caractéristiques des composants micro-ondes sont en règle générale mesurées à l'aide de signaux monochromatiques, dont la valeur instantanée est limitée à $\sqrt{2} V_{\text{eff}}$. Le fait que l'amplitude soit bornée constitue une différence fondamentale, en particulier le niveau de puissance correspondant à une erreur de linéarité donnée sera beaucoup plus faible dans le cas du bruit. Le niveau de saturation le plus souvent spécifié correspond à une erreur de 25 % (1 dB). C'est inacceptable si l'on veut pouvoir étalonner correctement les observations. De plus, une seule valeur comme le niveau de compression à 1 dB n'indique pas du tout comment le composant se comporte en dessous de ce niveau, si la saturation arrive de façon brusque ou graduelle. La seule méthode possible est donc de faire des mesures point par point et de travailler en dessous du niveau qui compromet la qualité de l'étalonnage.

Enfin, on veut que la densité spectrale des fluctuations du niveau de sortie V_s du détecteur soit aussi proche que possible de la valeur théorique V_s^2/B . On va voir plus loin que le fait que le signal de sortie soit proportionnel à la puissance d'entrée ne garantit pas forcément cette propriété.

3.3 Mesures de puissance en radiofréquence

Les systèmes de mesure de puissance utilisés en radiofréquence se réduisent essentiellement à deux types: les thermocouples et les détecteurs à diode. Ils sont basés sur des principes physiques très différents: les premiers sur des phénomènes thermoélectriques comme l'effet Seebeck et les seconds sur la nonlinéarité de la caractéristique tension courant des jonctions entre deux semiconducteurs de dopages différents ou entre un semiconducteur et un métal.

Les thermocouples ont l'avantage d'être bien linéaires, cependant leur constante de temps relativement longue due à leur principe même de fonctionnement rend impossible leur utilisation en mode de détection synchrone rapide. Leur sensibilité est assez faible (de l'ordre de $60 \mu\text{V}/\text{mW}$) ce qui implique l'emploi de puissances élevées si l'on ne veut pas être noyé dans le bruit. De plus, les dérives sont assez importantes et leur correction implique de fréquentes mesure du niveau de zéro.

Les détecteurs à diode, pour leur part, offrent une meilleure sensibilité et leur temps de réponse ne dépend quasiment que de la bande passante de l'amplificateur qui suit l'étage de détection. La dérive du niveau de zéro est donnée par la dérive de la tension d'erreur de l'amplificateur. Par contre la gamme de puissance sur laquelle le signal de sortie est proportionnel à la puissance incidente est limitée et les variations de sensibilité en fonction de la température de la diode sont importantes. Les modèles les plus couramment utilisés, du moins à large bande en micro-ondes, sont construits autour d'une diode Schottky dont nous allons voir plus loin qu'elle n'est pas exempte de défauts. On trouve par exemple à l'heure actuelle des modèles fonctionnant de 10 MHz à 50 GHz, et dont la réponse en fréquence est extrêmement plate.

Parmi les détecteurs à diode, il convient de faire la distinction entre ceux qui demandent que l'on injecte un courant de polarisation dans la diode et ceux qui n'en ont pas besoin. On peut d'emblée éliminer les premiers pour les applications qui demandent un couplage continu

du signal de sortie et une grande stabilité à cause du bruit et des dérives inévitablement introduits par le courant de polarisation.

Si l'on augmente le niveau de puissance d'entrée, le détecteur à diode Schottky devient un détecteur de niveau de crête du signal incident. Ce type de détecteur peut servir à mesurer ou stabiliser la puissance d'une onde monochromatique mais ne donne pas de résultat satisfaisant sur du bruit gaussien comme les signaux astronomiques. La technique utilisée sur les wattmètres de laboratoire qui consiste à effectuer une correction numérique du signal détecté par la diode pour augmenter la dynamique dépend fortement des hypothèses faites sur le signal d'entrée. Elle ne marche bien que sur des signaux monochromatiques: la présence d'harmoniques, même faibles, peut fausser la mesure qui dépend de la phase relative du fondamental et des harmoniques.

3.4 Terminologie des détecteurs à diode

Les circuits de détection micro-ondes se composent de deux parties:

- une partie radiofréquence de complexité variable, entre une simple liaison à la masse en continu et de complexes circuit d'adaptation d'impédance et de compensation pour les détecteurs à très hautes performances, cette partie se termine sur l'une des bornes de la diode.
- une partie vidéo, puisque c'est le terme consacré, qui part de l'autre borne de la diode et comprend au moins un condensateur d'intégration.

Les fabricants de diodes pour détecteurs et de détecteurs utilisent un jargon dont les principaux termes sont décrits ci-dessous. Les 3 paramètres les plus importants qui caractérisent un détecteur sont les suivants:

- Résistance video: c'est par définition la pente $\frac{dV}{dI}$ de la diode, à son point de fonctionnement, incluant le cas échéant le courant de polarisation, lorsque la puissance radiofréquence injectée est nulle.
- Sensibilité: exprimée en $\text{mV}/\mu\text{W}$, c'est le rapport entre la tension de sortie du détecteur et la puissance RF injectée. Elle est spécifiée sur une onde monochromatique dans des conditions optimales d'adaptation d'impédance de la diode, difficiles à réaliser dans le cas d'un circuit à très large bande. On peut aussi définir une sensibilité en courant en divisant l'efficacité en tension à faible niveau par la résistance vidéo.
- Puissance minimale détectable: (en anglais Tangential sensitivity) Elle dépend de la bande passante et du niveau de bruit de l'amplificateur vidéo utilisé. Ces deux derniers paramètres varient d'un fabricant à l'autre, et même d'un modèle à l'autre d'un même fabricant, ce qui oblige à un petit peu de gymnastique pour effectuer des comparaisons significatives. Les valeurs les plus courantes pour la bande passante s'échelonnent entre 1 et 10 MHz, la valeur de la sensibilité variant comme l'inverse de la racine carrée de cette bande, laquelle n'inclut jamais les fréquences inférieures à quelques dizaines de kilohertz, qui dégraderaient très fortement les chiffres annoncés à cause du bruit rose.

Les paramètres des diodes sont toujours spécifiés sous des conditions optimales d'adaptation d'impédance et à une fréquence donnée.

Ceux des détecteurs le sont souvent sous une impédance de $50\ \Omega$, valeur la plus courante de l'impédance caractéristique des composants et câbles micro-ondes. Comme l'impédance des diodes est en règle générale très supérieure à $50\ \Omega$, et que par contre la forme générale de leur caractéristique courant tension est la même, la puissance minimale détectable est plus forte dans le cas d'un détecteur que d'une diode.

3.5 Le détecteur à diode Schottky

Le but de ce paragraphe est de mettre à jour les limitations intrinsèques de la plupart des détecteurs à diode couramment disponibles.

Une diode Schottky est constituée par une jonction entre un métal et un semiconducteur de type N ou P. Contrairement aux diodes à jonction PN, il s'agit d'un composant à porteurs majoritaires qui n'exhibe pas de problème de temps de recombinaison et permet par conséquent son utilisation à des fréquences beaucoup plus élevées.

La caractéristique statique d'une diode Schottky est donnée par la formule suivante, dite de Richardson:

$$I(V) = I_s \left[\exp \left(\frac{q(V - I(V)R_s)}{nkT} \right) - 1 \right]$$

dans cette équation, I_s est une caractéristique de la diode appelée courant de saturation, R_s est la résistance du substrat qui diminue, par simple loi d'Ohm, la tension réellement appliquée à la jonction, k est la constante de Boltzmann, et T la température du composant. Quant à n , joliment appelé facteur d'idéalité de la diode; il est en général compris entre 1.05 et 1.25 à la température ambiante, à laquelle on peut prendre $nkT/q = 28\text{ mV}$ comme une bonne approximation. Signalons que I_s varie assez fortement avec la température et de façon négligeable en fonction de la tension aux bornes de la jonction. En particulier, le terme $e^{\frac{qV}{nkT}}$ diminue lorsque la température augmente mais la variation de I_s fait plus que compenser cet effet car son coefficient de température est positif et important ($\approx 1\ \%/K$ à 300 K). On peut en théorie calculer I_s en fonction des paramètres géométriques et physico-chimiques de la diode mais le résultat est très approximatif [31, page 18] et est absolument sans intérêt pour l'utilisateur de la diode.

Malgré ce qui vient d'être dit, les diodes Schottky obéissent bien à l'équation de Richardson. Dans le cas où l'on s'intéresse au fonctionnement en mode alternatif à haute fréquence, il faut tenir compte des éléments parasites de la jonction et du boîtier. Le modèle le plus couramment utilisé est montré en figure 3.1. En toute rigueur, la capacité C_j de la jonction dépend de la tension qui lui est appliquée mais cet effet est négligeable dans le cas d'un circuit fonctionnant à très bas niveau (moins de 200 mV pic à pic) comme un détecteur [31, page 27].

Les détecteurs coaxiaux commerciaux semblent tous réalisés selon un schéma proche de celui représenté à la figure 3.2. La seule variation importante concerne le circuit de mise à la masse en continu de l'entrée, il est soit réalisé avec une résistance égale à l'impédance caractéristique du système (50 ou 75 Ω), soit avec une inductance. Dans le premier cas, on obtient un faible taux de réflexion sur une très large bande; dans le second cas, la sensibilité est bien meilleure parce que la désadaptation provoque un doublement de la tension présentée à la jonction, mais la bande relative d'utilisation est plus étroite. La résistance de charge R_c , qui a pour but d'étendre la gamme sur laquelle la tension de sortie est proportionnelle à la puissance reçue, est quelquefois intégrée au détecteur mais plus souvent ajoutée par l'utilisateur pour optimiser sa valeur en fonction de l'application. La

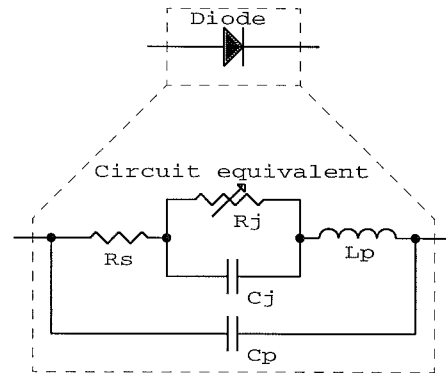


Figure 3.1: Modèle de diode micro-onde avec éléments parasites: C_j est la capacité de la jonction, R_s la résistance du substrat semiconducteur, L_p et C_p sont l'inductance de la « moustache » et la capacité du boîtier.

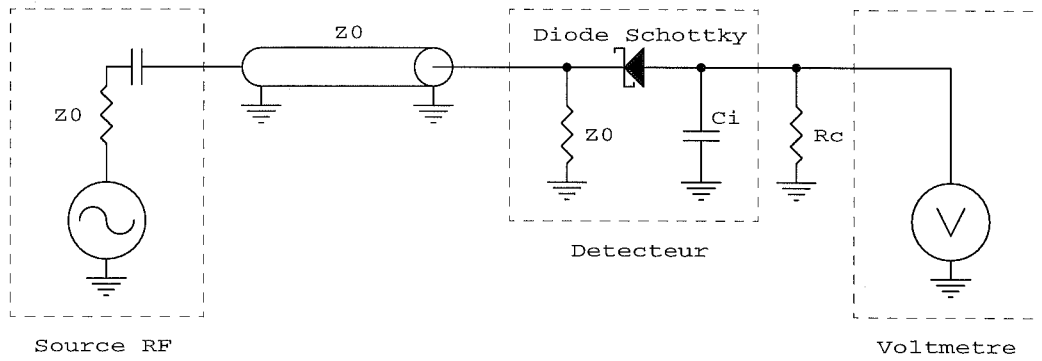


Figure 3.2: Schéma de détecteur coaxial dans son circuit d'application.

dispersion sur les caractéristiques des diodes est suffisamment importante pour que cette résistance doive être ajustée individuellement pour chaque détecteur.

3.5.1 Réponse en puissance du détecteur

Calculons maintenant la réponse du détecteur à diode Schottky dans le circuit représenté en figure 3.2 sous les hypothèses suivantes:

- On s'intéresse à des fréquences de sortie faibles, pour lesquelles la capacité d'intégration C_i est une haute impédance,
- Les fréquences composant le signal d'entrée sont suffisamment élevées pour que le condensateur C_i puisse être considéré comme un court-circuit,
- l'impédance de source du circuit est beaucoup plus faible que celle de la diode pour obtenir une réponse plate en fréquence, le circuit externe impose donc une tension à la jonction;
- les éléments parasites de la diode sont négligeables, on supposera que R_s , L_p et C_p sont nuls.

Les deux premières hypothèses sont parfaitement justifiées: la fréquence du signal d'entrée n'est pas inférieure à quelques dizaines de MHz et on s'intéresse au signal redressé jusqu'à quelques dizaines de kHz tout au plus. La troisième est vraie dans tous les cas de détecteurs non optimisés pour une fréquence spécifique (la fréquence de coupure $1/2\pi R_j C_j$ atteint au mieux 1.2 GHz); quant à la quatrième, le fait que l'on trouve des détecteurs commerciaux ayant une très faible ondulation sur plus de 20 GHz prouve qu'on peut, sinon négliger, du moins compenser les éléments parasites.

Evaluons donc la valeur moyenne pour une diode Schottky du courant redressé I_r lorsque le signal d'entrée V_e est du bruit gaussien:

$$I_r = \int_{-\infty}^{+\infty} I(V_e - V_s) P(V_e) dV_e$$

où V_s est la tension mesurée par le voltmètre de la figure 3.2, et avec:

$$I(V) = I_s \left[\exp \left(\frac{qV}{nkT} \right) - 1 \right]$$

$$P(V) = \frac{1}{\sqrt{2\pi} V_{\text{eff}}} \exp \left(-\frac{V^2}{2V_{\text{eff}}^2} \right)$$

$$V_s = I_r R_c$$

on obtient:

$$I_r = I_s \left[\exp \left(\frac{V_{\text{eff}}^2 q^2}{2n^2 k^2 T^2} - \frac{q I_r R_c}{nkT} \right) - 1 \right] \quad (3.3)$$

pour un signal sinusoïdal, dont la densité de probabilité est définie par: $P(V) = \frac{1}{\pi \sqrt{2V_{\text{eff}}^2 - V^2}}$, le résultat est différent:

$$I_r = I_s \left[I_0 \left(\frac{\sqrt{2} q V_{\text{eff}}}{nkT} \right) \exp \left(-\frac{q V_s}{nkT} \right) - 1 \right]$$

avec I_0 fonction de Bessel modifiée dont le développement en série est:

$$I_0(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2} \right)^{2k}}{(k!)^2}$$

Pour une résistance de charge infinie ($I_r = 0$), la fonction de transfert se réduit à:

$$V_s = \frac{q V_{\text{eff}}^2}{2nkT}, \text{ sur du bruit gaussien} \quad (3.4)$$

$$V_s = \frac{nkT}{q} \text{Log } I_0 \left(\frac{\sqrt{2} q V_{\text{eff}}}{nkT} \right), \text{ sur un signal monochromatique}$$

Sous une impédance de source de 50Ω , dans le circuit à adaptation résistive, cela donne pour une puissance incidente P_i :

$$V_s = \frac{q Z_0}{2nkT} P_i \approx 970 \mu\text{V} / \mu\text{W} \quad (3.5)$$

avec un coefficient de température non négligeable de 0.3 %/K en théorie, et expérimentalement plus proche de 1 %/K.

On pourrait croire à voir ce résultat que la diode Schottky est un détecteur idéal sur du bruit puisque la tension de sortie est rigoureusement proportionnelle à la puissance d'entrée. La différence avec le signal sinusoïdal vient du fait que la tension détectée ne saurait dépasser la valeur crête du signal d'entrée. Il faut toutefois rappeler ici les hypothèses sur lesquelles ce calcul repose et considérer qu'il n'est plus valable lorsque V_s atteint 2 ou 3 fois V_{eff} . Des mesures effectuées sur des détecteurs ont montré que la limite de validité de (3.5) est comprise entre 4 et 5 μW . On peut dans ce cas améliorer la linéarité en rajoutant au circuit la résistance de charge R_c au prix d'une diminution de la sensibilité. Il subsiste cependant un problème: quelles sont les fluctuations du niveau de sortie du détecteur ? C'est l'objet du paragraphe suivant.

3.5.2 Fluctuation du signal détecté

Rappelons tout d'abord le théorème selon lequel le spectre de puissance $S(\nu)$ d'un signal $V(t)$ dépendant du temps t est égal à la transformée de Fourier de sa fonction d'autocorrélation, notée $R(\tau)$ et définie par:

$$R(\tau) = \overline{V(t+\tau)V(t)}$$

cette fonction est maximale en 0 et on utilise fréquemment la fonction d'autocorrélation normalisée r_τ égale à:

$$r_\tau = \frac{R(\tau)}{R(0)}$$

Lorsque l'on considère deux signaux gaussiens V_1 et V_2 partiellement corrélés de taux de corrélation r , la densité de probabilité jointe a pour expression [30, page 129]:

$$P(V_1, V_2; r) = \frac{1}{2\pi V_{\text{eff}}^2 \sqrt{1-r^2}} \exp \left[\frac{-1}{2(1-r^2)} \left(\frac{V_1^2 - 2rV_1V_2 + V_2^2}{V_{\text{eff}}^2} \right) \right]$$

Calculons maintenant $R(\tau)$ du signal astronomique dans les conditions idéales d'un bruit gaussien de densité spectrale de puissance uniforme sur la bande de fréquences comprises entre f_i et f_s et zéro ailleurs, c'est simplement la transformée de Fourier de la somme de deux rectangles:

$$R(\tau) = \frac{\sin 2\pi f_s \tau - \sin 2\pi f_i \tau}{\pi \tau}$$

et après normalisation:

$$r_\tau = \frac{\sin 2\pi f_s \tau - \sin 2\pi f_i \tau}{2\pi(f_s - f_i)\tau}$$

Il faut ensuite calculer la fonction d'autocorrélation du courant redressé par la diode $R_{I_r}(\tau)$, qui est représentée par l'intégrale:

$$R_{I_r}(\tau) = \frac{I_s^2}{2\pi V_{\text{eff}}^2 \sqrt{1-r_\tau^2}} \iint \left(e^{\frac{q(V_1 - V_s)}{nkT}} - 1 \right) \left(e^{\frac{q(V_2 - V_s)}{nkT}} - 1 \right) P(V_1, V_2; r_\tau) dV_1 dV_2$$

Seul le terme produit des deux exponentielles exige d'effectuer explicitement les deux intégrations; les deux termes avec une exponentielle sont identiques et ne dépendent pas du taux de corrélation; quant au terme constant, il vaut 1 par définition de la densité de probabilité. Le calcul donne finalement:

$$R_{I_r}(\tau) = I_s^2 \left[\exp \left(\frac{q^2 V_{\text{eff}}^2 (1+r_\tau)}{n^2 k^2 T^2} - \frac{2qV_s}{nkT} \right) - 2 \exp \left(\frac{q^2 V_{\text{eff}}^2}{2n^2 k^2 T^2} - \frac{qV_s}{nkT} \right) + 1 \right]$$

et en utilisant (3.3)

$$R_{I_r}(\tau) = I_s^2 \left[\exp \left(\frac{q^2 V_{\text{eff}}^2 r_\tau}{n^2 k^2 T^2} \right) \left(1 + \frac{I_r}{I_s} \right)^2 - 2 \left(1 + \frac{I_r}{I_s} \right) + 1 \right]$$

Dans le cas de la résistance de charge infinie, cette formule se simplifie:

$$R_{I_r}(\tau) = I_s^2 \left[\exp \left(\frac{q^2 V_{\text{eff}}^2 r_\tau}{n^2 k^2 T^2} \right) - 1 \right]$$

Pour obtenir le spectre $S_{I_r}(\nu)$ du courant redressé par la diode, il faut maintenant, par application du théorème de Wiener-Khintchine, calculer la transformée de Fourier de $R_{I_r}(\tau)$. Remarquons tout d'abord que le seul terme qui dépende de τ dans la formule générale est à un coefficient multiplicatif près le même que dans la formule simplifiée. Les autres termes sont des constantes réelles, dont la transformée de Fourier est une simple distribution δ , qui représente la valeur moyenne du courant redressé qui passe dans la résistance de charge R_c , et qui est naturellement nul lorsqu'elle est infinie.

Evaluons $S_{I_r}(\nu)$ en développant l'exponentielle dans $R_{I_r}(\tau)$ et en posant $\alpha = \left(\frac{q V_{\text{eff}}}{n k T} \right)^2$:

$$R_{I_r}(\tau) = I_s^2 \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} r_\tau^k$$

$$S_{I_r}(\nu) = I_s^2 \int_{-\infty}^{+\infty} \exp(-2j\pi\nu\tau) \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} r_\tau^k d\tau$$

que l'on peut transformer en somme de transformées de Fourier en échangeant les signes somme et intégrales:

$$S_{I_r}(\nu) = I_s^2 \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} \int_{-\infty}^{+\infty} \exp(-2j\pi\nu\tau) r_\tau^k d\tau$$

la transformée de Fourier de r_τ est le spectre de puissance normalisé du signal d'entrée $s_{V_e}(\nu)$, et la transformée de Fourier de r_τ^n est la $k^{\text{ième}}$ puissance de convolution $s_{V_e}(\nu)$, notée $s_{V_e}(\nu)^{*(k)}$:

$$S_{I_r}(\nu) = I_s^2 \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} s_{V_e}(\nu)^{*(k)}$$

Or $s_{V_e}(\nu)$ est connu, pour du bruit idéalement plat et limité en bande centré en ν_0 et de largeur $\Delta\nu$, c'est:

$$s_{V_e}(\nu) = \frac{1}{2\Delta\nu} \left(\Pi \left(\frac{\nu - \nu_0}{\Delta\nu} \right) + \Pi \left(\frac{\nu + \nu_0}{\Delta\nu} \right) \right)$$

dans cette formule $\Pi(x)$ est la distribution porte valant 1 entre $-1/2$ et $+1/2$, et zéro ailleurs.

Dans le cas d'une bande relative très large, avec $2\Delta\nu$ à peine inférieur à ν_0 , mais lorsqu'on ne s'intéresse dans $S_{I_r}(\nu)$ qu'aux fréquences ν inférieures à $\nu_0 - \Delta\nu/2$, le terme correspondant à $k = 1$ dans $S_{I_r}(\nu)$ est nul et on peut faire l'approximation suivante pour les termes d'ordre $k \geq 2$:

$$s_{V_e}(\nu) = \frac{1}{2\Delta\nu} \Pi \left(\frac{\nu}{2\Delta\nu} \right)$$

le carré de convolution de $s_{V_e}(\nu)$ par lui même vaut:

$$s_{V_e}(\nu)^{(2)} = \frac{1}{2\Delta\nu} \Lambda\left(\frac{\nu}{2\Delta\nu}\right)$$

Λ est la distribution triangle définie par $\Lambda(x) = 1 - |x|$ pour $|x| < 1$ et zéro partout ailleurs. Tous les produits de convolution d'ordre supérieur ont des dérivées nulles en zéro et seul les termes d'ordres 3 et 4 seront nécessaires par la suite. On les calculera à l'origine seulement puisqu'on ne s'intéresse qu'à des fréquences proches de zéro et petites par rapport à $\Delta\nu$:

$$[s_{V_e}(\nu)^{(3)}]_0 = \int_{-\infty}^{+\infty} \frac{1}{2\Delta\nu} \Pi(x) \Lambda(-x) dx = \frac{3}{8\Delta\nu}$$

$$[s_{V_e}(\nu)^{(4)}]_0 = \int_{-\infty}^{+\infty} \frac{1}{2\Delta\nu} \Lambda(x) \Lambda(-x) dx = \frac{1}{3\Delta\nu}$$

Reportons maintenant ces termes dans l'expression de $S_{I_r}(\nu)$:

$$S_{I_r}(\nu) = I_s^2 \left[\frac{\alpha^2}{4\Delta\nu} \Lambda\left(\frac{\nu}{2\Delta\nu}\right) + \frac{\alpha^3}{16\Delta\nu} + \frac{\alpha^4}{72\Delta\nu} + \text{termes d'ordre supérieurs} \right]$$

aux fréquences auxquelles on s'intéresse, on pourra considérer que $\Lambda\left(\frac{\nu}{2\Delta\nu}\right) = 1$. Remarquons aussi que les termes d'ordre supérieurs sont tous strictement positifs, et que donc l'expression donne une minoration de la densité spectrale du courant redressé. Dans cette formule, le terme d'ordre 2 correspondrait à une diode parfaitement quadratique, et le spectre de puissance représente les fluctuations de puissance d'entrée. C'est le terme que l'on attend, on peut le vérifier simplement en calculant la tension équivalente dans le cas où la résistance de charge est absente:

$$s_{V_s}(\nu) = \frac{Z_v^2 S_{I_r}(\nu)}{V_s^2} = \frac{1}{\Delta\nu}$$

qui représente un bruit blanc de densité spectrale constante et qui donnera $1/\Delta\nu$ d'incertitude après une seconde d'intégration.

Le terme de troisième ordre (et ceux d'ordre supérieurs) représentent des fluctuations de la tension de sortie associés à des ordres supérieurs dans le développement en série de la caractéristique de la diode et donc à des moments d'ordres supérieurs dans la distribution de probabilité de la tension d'entrée. Ces termes sont indésirables puisqu'ils impliquent un accroissement de l'incertitude de la mesure qui apparaîtra sous la forme d'un bruit supérieur au bruit d'un détecteur quadratique parfait. Evaluons à quel niveau de puissance l'augmentation de temps d'intégration due au terme de troisième ordre atteint 10 %, il suffit de résoudre:

$$\frac{1}{10} \frac{\alpha^2}{4\Delta\nu} = \frac{\alpha^3}{16\Delta\nu}$$

et on obtient en substituant dans α sa définition

$$V_{\text{eff}}^2 = \frac{4n^2 k^2 T^2}{10q^2}$$

sous une impédance de 50Ω la puissance correspondante pour une diode à la température ambiante ($nkT/q = 28 \text{ mV}$) est environ $6 \mu\text{W}$. Par contre, si la bande de bruit est inférieure à une octave, le terme d'ordre 3 n'apparaît pas autour de la fréquence zéro. Dans ce cas, c'est le terme d'ordre 4, dont la valeur en zéro est indépendante de la bande utilisée, qui

donne la limite de puissance acceptable, de l'ordre de $45 \mu\text{W}$ dans les mêmes conditions que précédemment. Cette dernière valeur est une limite absolue, toutes conditions confondues, car l'augmentation de temps d'intégration croît très vite puisqu'elle est une fonction quadratique de la puissance. L'excès de bruit dû à l'intermodulation du troisième ordre a été observé expérimentalement sur un détecteur construit autour d'une diode Schottky (Hewlett-Packard HSCH-3486), et correspond bien à la prédiction théorique.

Il serait donc plus favorable de connecter les détecteurs sur la première fréquence intermédiaire des récepteurs, à 1.5 ou 4 Ghz, cette conclusion est indépendante du détecteur utilisé à partir du moment où la diode parfaitement quadratique n'existe pas. Le problème de l'ordre 3 n'existe pas dans les détecteurs des batteries de filtre de l'IRAM dont la fréquence centrale divisée par la largeur à mi-hauteur est toujours supérieure à 8.

3.5.3 Puissance minimale utilisable

Elle est évidemment donnée par le niveau de bruit de la diode, étant donné que densité spectrale du signal issu d'un détecteur quadratique parfait vaut V_s^2/B et que le bruit propre de la diode Schottky à faible niveau est égal à celui d'une résistance de valeur égale à son impédance vidéo $Z_v = nkT/qI_s$, qui a pour densité spectrale $4kTZ_v$, le critère des 10 % donne

$$V_s^2 > \frac{40nk^2T^2B}{qI_s}$$

et dans le cas où on peut appliquer (3.5)

$$P_i > \frac{4\sqrt{10}nkT}{Z_0} \frac{1}{q} \sqrt{kTBZ_v}$$

qui donne environ $1 \mu\text{W}$ pour $Z_0 = 50 \Omega$, $B = 1 \text{ Ghz}$ et $Z_v = 5 \text{ k}\Omega$ (cas de la diode HSCH-3486).

Cette valeur minimale pêche évidemment par optimisme, étant donné que nous n'avons tenu compte jusqu'ici ni du bruit du circuit amplificateur qui suit le détecteur, ni surtout des problèmes de bruit rose.

Le bruit rose est un phénomène très général, qui affecte la plupart des phénomènes physiques et, à des degrés divers, tous les composants électroniques. Il est appelé ainsi parce que sa densité spectrale, au lieu d'être indépendante de la fréquence ν comme par exemple le bruit thermique, est caractérisé par une dépendance en ν^α avec $\alpha \approx -1$. L'origine de ce bruit n'est toujours pas bien comprise et il n'existe pas de théorie générale permettant de le décrire, bien qu'un certain nombre de modèles disparates donnent des résultats satisfaisants dans des cas particuliers [32, chapitre 8]. La minimisation de cette source de bruit dans un composant ou dans un circuit est donc plus souvent une question d'expérience et d'empirisme que de connaissances théoriques.

Dans l'immense majorité des cas, on peut considérer que le bruit d'un système électronique est la somme d'une composante de bruit rose et d'une de bruit blanc, que l'on peut caractériser, comme nous le ferons par la suite, par la densité spectrale S_∞ dans la limite des hautes fréquences et la fréquence ν_r , dite de remontée du bruit rose, à laquelle les deux contributions sont égales. Dans ce cas, la forme du spectre de bruit s'écrit simplement

$$S(\nu) = S_\infty(1 + \nu_r/\nu) \quad (3.6)$$

Si un bruit tel bruit passe dans un filtre parfait sélectionnant les fréquences comprises entre ν_0 et ν_1 , la valeur quadratique moyenne S du signal filtré est

$$S = \int_{\nu_0}^{\nu_1} S(\nu) d\nu = S_{\infty} (\nu_1 - \nu_0 + \nu_r \text{Log}(\nu_1/\nu_0)) \quad (3.7)$$

On voit ici que la bruit rose donne une contribution constante par octave, et non proportionnel à la largeur de bande comme le bruit blanc.

3.5.4 Conséquences

La première conséquence à tirer de ce qui précède est qu'il ne suffit pas d'optimiser le détecteur pour obtenir une réponse linéaire en puissance, il faut aussi vérifier que les fluctuations du signal détecté correspondent à la valeur attendue. En fait on peut dire que les deux mesures sont indépendantes, cela ne ressort pas de façon limpide de la littérature, et le contre-exemple correspond précisément au modèle théorique du type de diode le plus utilisé à l'heure actuelle. L'autre inconvénient de la nécessité des deux mesures est que la deuxième est nettement plus difficile à réaliser en pratique: elle demande une source de bruit dont la stabilité à basse fréquence soit connue et un instrument capable d'effectuer une analyse spectrale du signal détecté.

La deuxième conséquence est que la diode Schottky est loin d'être un détecteur parfait et qu'on ne peut pas jouer impunément sur la résistance de charge R_c , qui permet assez facilement de rendre la diode linéaire jusqu'à des puissances de plusieurs dizaines de microwatts, pour travailler avec des signaux plus forts et éliminer les problèmes de bruit, sans détériorer la bande effective d'intégration, qui est un paramètre fondamental en radio-astronomie.

On pourrait essayer d'augmenter la dynamique de deux façons, la première consiste à contrôler la température de la diode pour optimiser ses caractéristiques, mais il suffit de reprendre les expressions des paragraphes précédents pour constater que les puissance minimale et maximale acceptables varient toutes les deux comme le carré de la température. La deuxième méthode consiste à utiliser plusieurs diodes en réseau série-parallèle, on peut facilement voir dans ce dernier cas que, sous l'hypothèse optimiste de diodes rigoureusement identiques, la dynamique augmente comme la racine carrée du nombre diodes, quelle que soit la configuration du réseau. On en arrive donc vite à un circuit monstrueux, sans parler des problèmes des éléments parasites, et la seule solution pratique serait d'intégrer les diodes.

3.6 Diodes tunnel

Comme on le voit, la diode Schottky, avec sa caractéristique exponentielle, souffre de sérieuses limitations pour des bandes très supérieures à l'octave. Il fallait pourtant bien résoudre le problème pratique consistant à réaliser un détecteur ayant des performances satisfaisantes pour le télescope. Il existe un autre type de composant non-linéaire utilisé en détection: les diodes inverses (« backward » en anglais) qui sont, d'après la littérature, une variation sur le thème de la diode tunnel avec une région de résistance dynamique négative moins marquée ou même absente. C'est ce type de composant qui est utilisé dans les batteries de filtres de l'IRAM, mais la capacité parasite du modèle utilisé (BD4) empêche son emploi à des fréquences dépassant quelques centaines de mégahertz. Du point de vue physique, ce qui caractérise ces diodes est un taux de dopage extrêmement élevé, à la limite du niveau de solubilité des impuretés dans le cristal. Ceci signifie que leur résistance vidéo,

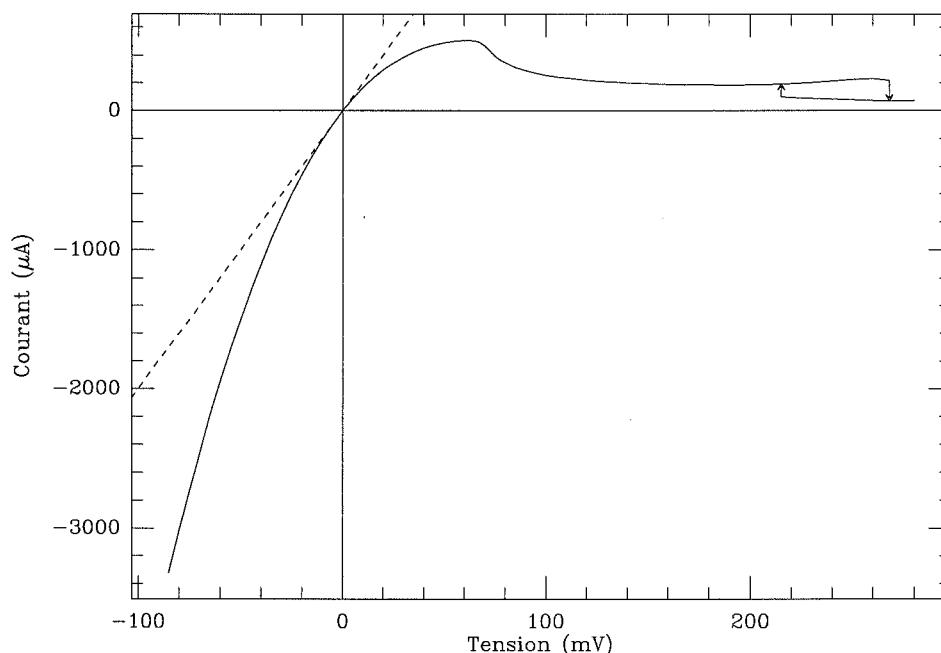


Figure 3.3: Caractéristique d'une diode tunnel de détecteur (type 23525HFBD). La ligne pointillée représente une résistance de $50\ \Omega$.

même sans courant de polarisation, sera beaucoup plus faible qu'une diode Schottky et donc leur niveau de bruit sera plus bas. De plus, et comme Esaki, leur inventeur, le remarquait déjà en 1958, les diodes tunnel possèdent un bruit rose extrêmement faible, non mesurable pour des fréquences supérieures à 10 Hz [32, page 156].

Il se trouve que ce type de diode est tombé en désuétude et que la plupart des fabricants de semiconducteurs qui les proposaient dans leur catalogue ont soit abandonné leur production, soit disparu. Ceci a posé de stupides, mais réels, problèmes d'approvisionnement au moment de l'étude de ce détecteur. Finalement, le processus de sélection, basé sur le seul critère du bruit, a abouti au choix du modèle 23525HFBD de Custom Components Inc., présenté dans un boîtier pour applications en ligne microruban (microstrip). Il ne s'agit pas là d'un composant de grande série: chaque diode est livrée dans une pochette plastique individuelle, elle même insérée dans une enveloppe sur laquelle sont marqués un numéro de série avec la date de fabrication, la capacité parasite et le courant de pic.

La caractéristique d'une de ces diodes est montrée en figure 3.3. La région de résistance dynamique négative y est très nettement marquée et caractéristique de l'effet tunnel. La discontinuité avec hystérésis qui apparaît entre 220 et 250 mV est située au delà de la zone de travail du détecteur.

C'est apparemment la diode avec la plus faible résistance vidéo disponible commercialement: $Z_v = 50\ \Omega$. Cette valeur est égale à l'impédance caractéristique des câbles et lignes micro-ondes et le taux d'onde stationnaire présenté par la diode pour de petits signaux sera donc proche de 1. Cette impédance correspond à un courant de pic de l'ordre de $500\ \mu\text{A}$. La capacité parasite totale de la diode encapsulée est de l'ordre de 0.6 à 0.7 pF. D'après le fabricant, la capacité de la jonction C_j ne dépasse pas 0.25 pF et la résistance série R_s est environ $7\ \Omega$, ce qui donne une constante de temps $\tau = R_s C_j \approx 1.6\ \text{ps}$, permettant d'envisager

leur utilisation à des fréquences de plusieurs dizaines de GHz.

Au point de vue de l'utilisateur, ces composants se distinguent par une certaine fragilité, et ce pour deux raisons: une tension de quelques centaines de millivolts induit un courant suffisamment fort pour les détruire, et le matériau de base est un cristal de germanium qui ne supporte pas les hautes températures, souder la diode sans précaution peut par exemple se traduire par une dégradation irréversible de ses caractéristiques.

On pourrait essayer d'effectuer sur ces composants une analyse comparable à celle réalisée sur les diodes Schottky, il faut pour cela disposer du développement de la caractéristique de la diode. Il se trouve que la comparaison de la caractéristique obtenue avec l'expression donnée par [33, equation (10.55)] ne donne pas de résultat satisfaisant. Si l'on ne considère que les tensions inférieures au pic de courant, une simple parabole donne un meilleur accord; par contre, aucun polynôme de degré raisonnable n'arrive à reproduire correctement la caractéristique si l'on tient compte de la région de résistance dynamique négative. Il est donc clair que, pour un signal suffisamment faible, le problème de l'intermodulation d'ordre 3 sera moindre, mais la limite supérieure sur le niveau de puissance acceptable sera déterminée expérimentalement.

La réalisation d'un circuit prototype a clairement montré que la valeur de la résistance de charge R_c qui donne la meilleure linéarité est nulle; autrement dit il faut mesurer le courant et non la tension détectée par la diode, ce qui se réalise simplement avec un amplificateur opérationnel dont l'entrée inverseuse constitue une masse virtuelle. Ce point, qui contredit la documentation des batteries de filtres de l'IRAM, est dû au fait que l'on travaille sous impédance adaptée. Pour le voir, il suffit de se rendre compte que lorsqu'on effectue une détection en tension, la diode se contre-polarise, ce qui augmente son impédance et donc le taux de réflexion et la tension présentée à la jonction. La tension détectée croît alors plus vite que la puissance disponible. Le rendement en détection de ces diodes est excellent, entre 10 et 12 $\mu\text{A}/\mu\text{W}$, à comparer aux 20 $\mu\text{A}/\mu\text{W}$ qui représentent le rendement théorique maximal et complètement irréalisable d'une diode Schottky.

La valeur minimale de l'admittance dynamique déterminée à partir de la courbe de plusieurs diodes est environ $-1/70\Omega^{-1}$, et ne pose donc en théorie aucun problème de stabilité dans un système 50Ω parfait. Cette valeur est cependant suffisamment proche de la limite pour demander au moins une vérification expérimentale de l'absence d'oscillations non contrôlées lorsqu'on applique un signal suffisamment fort.

3.7 Contrôle thermique

Les caractéristiques des semiconducteurs varient fortement en fonction de la température. La technique habituellement utilisée pour minimiser le problème dans le cas d'un amplificateur par exemple est de contrôler le gain par des composants plus stables qui constituent un réseau de contre-réaction.

Il n'existe aucune méthode comparable permettant de contrôler le coefficient de conversion entre puissance et tension ou courant d'un détecteur. On a signalé plus haut que le rendement d'un détecteur à diode Schottky varie d'environ 1%/K; ce chiffre est beaucoup plus faible pour les diodes tunnel, d'un facteur proche de 20.

Malgré cela, et peut-être parce que le coefficient de température des détecteurs à diode Schottky initialement prévus est assez effrayant, il a été décidé d'installer les détecteurs dans une enceinte stabilisée en température. Le principe du circuit de régulation, de type linéaire, est simple: un transistor de puissance chauffe l'enceinte et compense les fuites thermiques.

Un capteur (un simple transistor) permet de comparer la température à une consigne et ajuste le courant du transistor de chauffage en conséquence. Etant données la difficulté de modéliser correctement les phénomènes thermiques et d'évaluer la valeur des grandes constantes de temps associées, la mise au point de ce système, réalisé par H. Houeix et G. Butin à Grenade, a demandé beaucoup de patience pour éviter les oscillations et donner la meilleure régulation possible. L'ensemble est un petit peu mieux protégé qu'une centrale nucléaire contre les surchauffes éventuelles puisqu'on compte pas moins de cinq systèmes indépendants de limitation de température et de courant de chauffage.

La stabilité obtenue est de l'ordre du centième de degré sur une table de laboratoire, elle doit être encore meilleure une fois que les détecteurs sont montés dans la boîte de distribution, qui représente une enceinte thermique supplémentaire, elle-même installée dans une salle climatisée.

La température à laquelle travaillent les détecteurs est très stable mais n'est pas connue avec précision: environ 55°C. Cette valeur assez élevée a été choisie expérimentalement pour assurer que le courant dans le transistor de puissance ne tombe jamais à zéro. Elle reste suffisamment basse pour ne pas compromettre la fiabilité, mais contribue à augmenter légèrement le bruit propre du détecteur dont on va voir dans le prochain paragraphe qu'il est très proche de la limite thermique.

3.8 Amplification

Le signal qui sort du détecteur est à un niveau très faible, qui doit être fortement amplifié avant tout traitement ultérieur. Comme on exige une très grande dynamique à cause du facteur $1/B$ (120 dB de dynamique vidéo dans un hertz de bande pour 10 dB de dynamique astronomique utilisable) les considérations de bruit jouent un rôle essentiel et dictent le choix des composants, surtout celui du premier amplificateur opérationnel.

Dans le modèle le plus couramment utilisé, le bruit d'un amplificateur opérationnel est donné par 4 paramètres qui sont les valeurs de son bruit propre en tension e_n et en courant i_n , et les fréquences de remontée du bruit rose associées ν_{re} et ν_{ri} .

En l'absence de signal appliqué au détecteur, la diode peut être considérée équivalente à une résistance source égale à son impédance vidéo, et dans ce cas le bruit équivalent en tension V_n à l'entrée de l'amplificateur et la fréquence de remontée ν_{rv} du bruit rose correspondant valent:

$$V_n = \sqrt{4kTZ_v + Z_v^2 i_n^2 + e_n^2}$$

$$\nu_{rv} = \frac{Z_v^2 i_n^2 \nu_{ri} + e_n^2 \nu_{re}}{4kTZ_v + Z_v^2 i_n^2 + e_n^2}$$

Armé de ces formules, on peut comparer le résultat théorique obtenu pour le modèle classique d'amplificateur à faible bruit (OP27, proposé par plusieurs fabricants), et celui d'un modèle plus récent, le LT1028 de la société Linear Technology. En fait ces deux amplificateurs sont très proches dans leur conception, la seule solution pour diminuer le bruit d'un transistor bipolaire consistant à le faire travailler avec un courant de collecteur relativement fort, presque 1 mA pour le LT1028. Dans les deux cas, il s'agit d'amplificateurs avec une tension d'erreur faible et peu sensible à la température.

En utilisant les paramètres donnés par les fabricants présentés dans la table 3.1, V_n et ν_{rv} valent respectivement 3.15 nV.Hz^{-1/2} et 2.5 Hz pour l'OP27, et 1.25 nV.Hz^{-1/2} et 2.0 Hz

Modèle	e_n (nV.Hz ^{-1/2})	ν_{re} (Hz)	i_n (pA.Hz ^{-1/2})	ν_{ri} (Hz)
OP27	3.0	2.7	4	140
LT1028	0.85	3.5	9	250

Table 3.1: Caractéristiques de bruit des amplificateur opérationnels OP27 et LT1028 à 300 K.

pour le LT1028. La différence est considérable puisque le LT1028 améliore la dynamique du détecteur de 4 dB vers les hautes fréquences et 4.5 dB dans la limite des très basses fréquences où l'équation 3.6 montre que le critère à utiliser est $V_n\sqrt{\nu_{rv}}$. Un électronicien doublerait ces chiffres puisqu'il les prendrait à l'entrée de l'amplificateur et non du détecteur. Remarquons que, lorsqu'on néglige le bruit rose, le bruit de la diode et celui de l'amplificateur sont presque égaux; cela signifie que l'on ne peut réduire le bruit de façon significative qu'en améliorant les deux composants: cette situation est considérée comme optimale par les spécialistes.

Le calcul de bruit présenté précédemment est légèrement simplifié, car, dans le cas de ces amplificateurs à faible bruit qui disposent d'un circuit de compensation du courant de base des transistors d'entrée, le bruit en courant devrait en toute rigueur être séparé entre le bruit corrélé entre les deux entrées et le bruit non corrélé. Le bruit corrélé peut être éliminé en s'assurant que les impédances de source des deux entrées sont égales, ce qui s'obtient facilement en rajoutant une résistance entre l'entrée non inverseuse de l'amplificateur et la masse. La contribution au bruit de cette résistance serait cependant supérieure à la réduction de bruit attendue: les courbes fournies par le fabricant le montrent clairement, l'entrée non inverseuse a donc été délibérément connectée directement à la masse.

Quant aux résistances et condensateurs qui déterminent le gain et la bande passante du détecteur, ils ont été choisis de façon à obtenir en deux étapes un gain compris entre 350 et 400, ce qui donne à la sortie une pleine échelle proche de 7 V et le temps de montée de chaque étape a été fixé un peu arbitrairement à 10 μ s. La bande passante résultante est proche de 10 kHz.

Signalons que si ce nouvel amplificateur n'avait pas existé, il aurait été possible d'obtenir des caractéristiques de bruit similaires, et peut être même un petit peu meilleures, en utilisant des paires de transistors appariées. La mise en œuvre de ces composants est cependant nettement plus délicate que celle d'un simple amplificateur opérationnel et ne justifie pas l'effort nécessaire lorsque l'amélioration attendue ne dépasse pas 10 % dans l'hypothèse la plus optimiste.

Enfin, et comme la documentation de l'amplificateur permettait de penser qu'il pouvait y avoir une dispersion importante de ses caractéristiques de bruit, un tri systématique consistant à mesurer la densité spectrale de bruit à 10 Hz dans une bande de 1 Hz a été effectué. Il faut ici remercier l'un des opérateurs du télescope, J. Ranz, qui a aidé à réaliser cette opération, fastidieuse puisqu'elle demande près d'une heure par amplificateur: après avoir mis le circuit sous tension, il faut le laisser se stabiliser pendant au moins une demi-heure avant de pouvoir effectuer une mesure valable qui dure 20 minutes pour une incertitude statistique de 3 %. La conclusion qu'on peut en tirer est que cela n'a quasiment servi à rien, la dispersion sur les 50 amplificateurs mesurés n'étant pas significative. Un autre lot de LT1028 aurait peu être donné des résultats différents, mais il semble bien que les limites données par le fabricant soient surtout destinées à pouvoir les garantir à 100 % en ne réalisant qu'un test statistique de chaque lot.

3.9 Réalisation pratique

Le détecteur en lui même est un circuit assez particulier, combinant une partie radiofréquence d'un côté et basse fréquence à couplage continu et très faible bruit de l'autre. Le dessin définitif en est montré page 47. Chaque circuit est placé dans une alvéole individuelle fraisée, la cinquième alvéole centrale reçoit les transistors de chauffage et de mesure de température ainsi qu'un thermocontact de sécurité. Le détecteur est implanté sur une carte imprimée de dimensions réduites (33×86 mm), plutôt étroit pour minimiser la largeur totale de la boîte, seul paramètre important d'encombrement.

Le schéma utilisé se singularise par quelques détails notables, essentiellement parce que la cathode de la diode est à la masse. L'anode voit donc à la fois la partie radiofréquence et vidéo et se trouve placée dans une structure ressemblant à un diplexeur. Cette architecture fait économiser une capacité et une inductance par rapport aux circuit classique, puisque c'est le même condensateur qui isole d'une éventuelle composante continue et qui sert d'intégrateur haute fréquence, et que la bobine de mise à la masse en continu devient un composant du filtre passe-bas du courant détecté. La suppression de deux composants réactifs, source d'imperfections, est toujours appréciable dans un circuit micro-ondes, surtout s'il doit fonctionner sur plusieurs octaves.

Dans tout circuit large bande, les défauts des composants doivent être prises en compte. C'est ici le cas du filtre d'entrée qui doit rejeter sans résonance parasite toutes les fréquences comprises entre 80 et 1100 MHz et ne laisser passer que les fréquences inférieures à quelques centaines de kilohertz vers le premier amplificateur, qui manifeste « bruyamment » son mécontentement en cas de présence de signaux radiofréquence sur son entrée. Ce filtre se trouve de fait dans une position bizarre et sert à la fois de passe haut pour le signal radiofréquence et de passe-bas pour l'amplificateur. Il est attaqué sous une impédance de 50Ω à basse fréquence et de 25Ω à haute fréquence, et sa résistance de charge est nulle. La dernière inductance doit de plus être amortie par une résistance en parallèle pour éviter des oscillations de l'amplificateur. Toutes ces particularités, plus le fait qu'il faut éviter les résonances sur plus de trois décades au delà de la fréquence de coupure, font que ce filtre ne correspond pas aux schémas classiques, comme Butterworth ou Tchebyscheff, qui cherchent en général à limiter la dispersion sur les valeurs des composants, et ne considèrent que les réseaux constitués exclusivement d'éléments réactifs. En pratique, ce sont surtout les inductances dont le comportement devient vite imprévisible et pire, non reproductible.

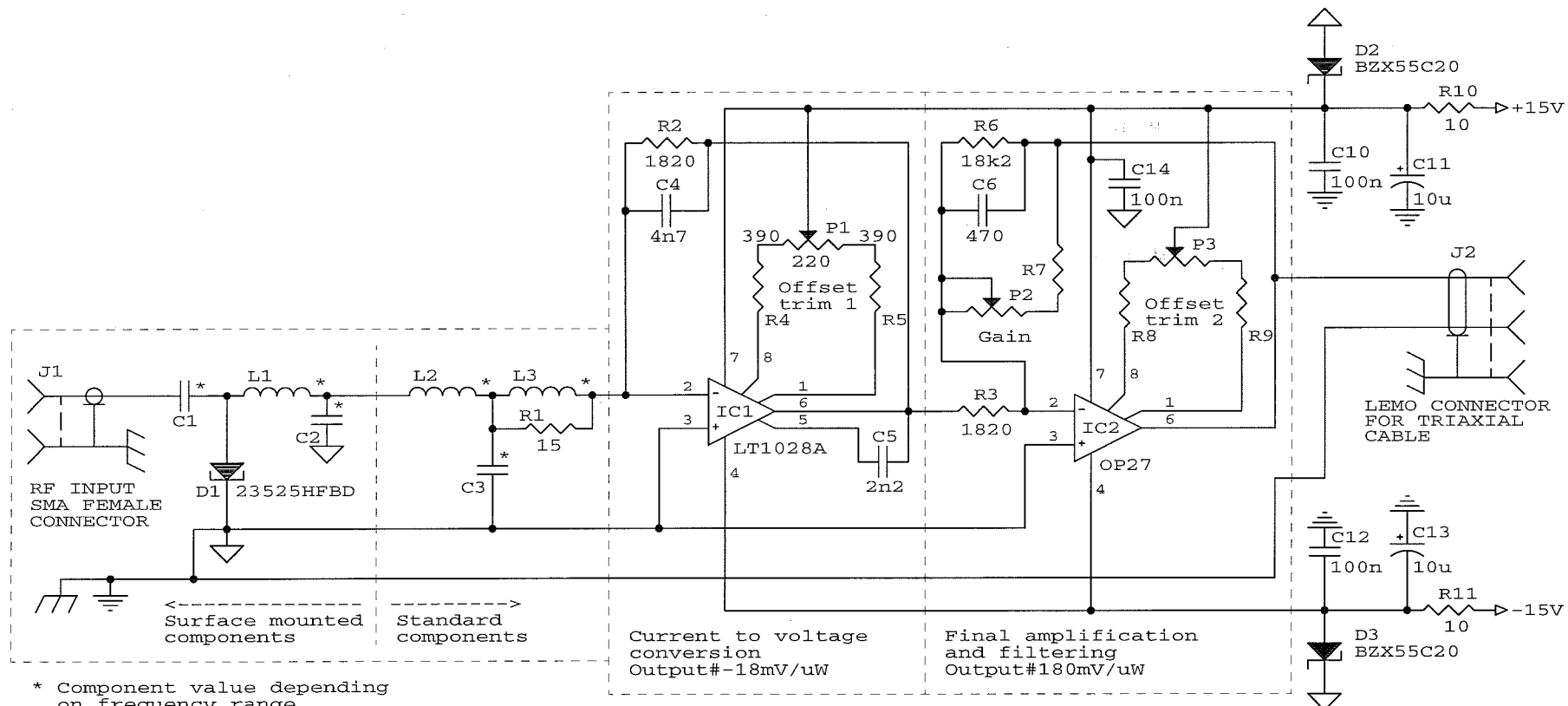
Au lieu de se lancer dans des calculs sans fin, les valeurs, limitées aux standards, des composants ont été déterminées après simulation du circuit sur ordinateur en assimilant la diode à une résistance de 50Ω . Cela demande d'avoir un peu d'expérience pour donner des valeurs initiales raisonnables aux composants mais converge assez rapidement.

Les valeurs indiquées sur le schéma ont ensuite été adoptées après vérification sur une carte prototype. La courbe de réponse théorique est plate dans ± 0.05 dB dans la bande du récepteur et possède deux pics de réjection dont un proche de 10 MHz; ce zéro de l'impédance a été placé là parce que c'est l'une des fréquences de référence du télescope qui a tendance à un peu trop se « promener » sur les câbles de fréquence intermédiaire.

Etant donné le niveau de bruit et la stabilité requises, le circuit de masse exige une attention particulière. Dans le détecteur, il est fondamental d'assurer que le point de masse de la diode et les entrées non inverseuses des amplificateurs soient rigoureusement au même potentiel, ce qui est réalisé par un circuit en étoile autour de la cathode de la diode.

Il reste sur le schéma deux diodes entre les alimentations et la masse dont le rôle n'a pas été expliqué jusqu'ici. En cas de défaillance de l'une des deux alimentations, et comme la

Figure 3.4: Schéma du détecteur



* Component value depending on frequency range.
For 80 Mhz to 1200 Mhz the values are:
C1=330pF chip
C2=1nF chip
C3=3.3nF plastic
L1=330nH with the pin on the diode side wound 2 turns of 3mm diameter
L2=1000nH
L3=3300nH

Cable R6 for fixed gain,
R7 and P2 for adjustable gain.
All resistors except R1, R10 and R11 are
1% tolerance with low temperature coefficients.

Notes:

- /// is case ground
≡ is power ground
▽ is signal ground

Instituto de Radioastronomia Milimetrica			
Title			
CONTINUUM DETECTORS			
Size	Document Number		REV
A	GP-8910		A
Date:	September 4, 1990	Sheet	1 of 1

diode tunnel se trouve directement entre la masse et l'une des entrées du premier amplificateur, ce dernier se retrouve alimenté par ses entrées; dans ce cas, le courant qui traverse la diode la détruit, l'amplificateur devient vite brûlant et ses caractéristiques de bruit se trouvent irrémédiablement dégradées. En présence des deux diodes (qui sont des zeners pour essayer aussi de protéger en même temps les amplificateurs en cas de surtension, l'idéal serait d'utiliser des supprimeurs de transitoires), l'alimentation manquante se retrouve à 0.7 V près au potentiel de la masse. Cela suffit à éviter les deux problèmes précédents. Depuis que cette protection a été ajoutée, plus aucun cas de destruction de la diode tunnel n'est à signaler. Pour l'anecdote, signalons que ce problème n'a été compris que le jour où un fil d'alimentation a été accidentellement arraché et où les signaux de fumée émis par l'amplificateur ont indiqué le problème, jusque là, un certain nombre de diodes étaient mortes mystérieusement, certainement à cause d'un retard d'une des alimentations au moment de la mise sous tension.

3.10 Acquisition

Le signal amplifié de la diode doit finalement être numérisé et fourni à l'ordinateur. La meilleure méthode pour mesurer la puissance moyenne reçue pendant un intervalle de temps donné est d'intégrer la tension de sortie du détecteur pendant cet intervalle et à numériser la valeur de l'intégrale.

3.10.1 Considérations générales

Les durées d'intégration couramment utilisées en continuum dans le domaine millimétrique sont en général comprises entre 10 ms et 1 s, et ne posent par conséquent pas de problèmes de rapidité. Ce sont donc essentiellement les caractéristiques de linéarité et de résolution qui influencent les choix techniques.

En ce qui concerne la linéarité, et étant donné ce qui a été dit au sujet des méthodes d'observation, il est clair que le problème sera essentiellement lié aux erreurs de linéarité pour de petites variations de la puissance d'entrée, autrement dit ce que l'on appelle la linéarité différentielle, qui peut être assez mauvaise dans le cas des convertisseurs numériques analogiques à approximation successive (voir le chapitre 4 et une astuce pour la diminuer).

Le signal de l'ancien détecteur continuum était numérisé par un convertisseur tension fréquence (un énorme et coûteux circuit hybride). C'est la même méthode qui a été choisie pour des raisons de simplicité (il n'y avait pas à reconcevoir l'interface avec l'ordinateur) et de performances.

3.10.2 Conversion tension-fréquence

Fonctionnement

Un convertisseur tension-fréquence se compose essentiellement de deux parties:

- un intégrateur, qui charge un condensateur avec un courant proportionnel à la tension d'entrée,
- et d'un circuit qui, chaque fois que la charge du condensateur atteint un seuil de déclenchement, enlève du condensateur une charge fixe.

Il suffit alors de compter le nombre de déclenchements du circuit de décharge du condensateur entre les instants t_0 et t_1 pour obtenir une mesure de l'intégrale $\int_{t_0}^{t_1} V(t) dt$. L'interface avec un ordinateur est donc simple puisqu'il suffit d'un compteur d'impulsions. L'avantage de ce genre de circuit est que, étant construit autour d'un intégrateur, il n'y a pas de problème de fréquence d'échantillonnage minimale en fonction de la bande passante du détecteur et sa linéarité est excellente.

Dans la plupart des convertisseurs tension-fréquence, la génération des impulsions de décharge est réalisée par un monostable qui commute une source de courant. La constante de temps du monostable est alors donnée par un simple réseau résistance capacité. Il est clair que dans ce cas, la fréquence maximale de sortie est toujours inférieure à l'inverse de la durée des impulsions, souvent d'un bon facteur 2. Le plus gros problème associé à ce type de circuits est la stabilité de la constante de temps: celle-ci varie en fonction de la température et aussi en fonction de l'intervalle entre deux impulsions. Ces deux phénomènes sont essentiellement liés au condensateur du monostable: le premier à la variation de la capacité avec la température et le deuxième à l'absorption lente de charges par le diélectrique. Ceci oblige à choisir des condensateurs de très haute qualité.

Le convertisseur utilisé ici élimine ces problèmes en synchronisant les impulsions de décharge du condensateur d'intégration sur une horloge externe. Son fonctionnement interne est un peu plus compliqué mais la stabilité et la linéarité s'en trouvent grandement améliorées, et ce sans aucun composant vraiment critique. Seul le condensateur d'intégration doit être à faible fuite, mais ni sa valeur exacte, ni sa stabilité en température n'entrent en compte puisque le principe de base du système est d'équilibrer la charge injectée par l'intégrateur par celle extraite par le générateur d'impulsions.

La fréquence maximale f_{\max} de sortie est égale à la moitié de la fréquence d'horloge, qui a été fixée à 2 Mhz.

Propriétés

Contrairement à un convertisseur analogique-digital, la résolution dépend ici du temps d'intégration. En première approximation, on peut considérer que si l'on compte les impulsions comprises entre les instants t_0 et t_1 , la résolution équivalente sera égale au nombre maximum d'impulsions soit $f_{\max}(t_1 - t_0)$ ce qui donne entre 10000 points pour 10 ms d'intégration et 10^6 (20 bits) pour une seconde. Ces chiffres, surtout le dernier, peuvent paraître optimistes, mais toutes les mesures effectuées sur ces convertisseurs montrent qu'on peut estimer leur résolution effective comme au moins égale à 22 bits.

Quant à l'incertitude statistique de la mesure, ou bruit de quantification, pour le calculer il faut savoir que le système d'acquisition dans sa forme présente est asynchrone, c'est à dire que les impulsions du convertisseur n'ont aucune relation de phase avec les instants t_0 et t_1 . Supposons pour le moment que la fréquence des impulsions est fixe, leur nombre entre t_0 et t_1 vaudra, en notant $y = f(t_1 - t_0)$, soit $E(y)$ soit $E(y) + 1$ avec pour probabilités respectives $E(y) + 1 - y$ et $y - E(y)$. A l'inverse cela signifie que pour si l'on compte n impulsions, la distribution de probabilité de y est:

$$P(y) = \Lambda(y - n)$$

donc la variance sur la mesure introduite par la quantification vaudra:

$$\sigma^2(y) = \int (y - n)^2 P(y) dy = \frac{1}{6}$$

pour un convertisseur analogique-numérique, on obtiendrait au contraire 1/12.

Avec ce résultat, on peut maintenant calculer la dynamique du système d'acquisition si l'on ne veut pas que la contribution de cette incertitude dépasse 10 %. Pour une bande de fréquence de largeur $\Delta\nu$, une durée d'intégration t et un nombre d'impulsions n , cette condition devient:

$$\frac{n^2}{\Delta\nu t} > \frac{5}{3}$$

ce qui signifie que la fréquence minimale d'impulsion f_{\min} doit être:

$$f_{\min} = \sqrt{5\Delta\nu/3t}$$

Sachant que f_{\max} vaut 1 MHz, on obtient donc une dynamique utilisable d pour $\Delta\nu = 1$ GHz:

$$d = \sqrt{600t}$$

qui donne une dynamique qui varie de 2.5 pour 10 ms à 25 pour 1 seconde d'intégration. La première valeur est franchement mauvaise, et c'est ce qui limite la dynamique utilisable des détecteurs. Ce n'est pas très grave pour l'instant à l'IRAM tant que la bande des récepteurs ne dépasse pas 500 MHz et que le temps d'intégration le plus court utilisé en pratique est 35 ms avec le « beam-switch », donnant une dynamique proche de 6. Pour les mesures demandant la plus grande sensibilité, on utilise le secondaire oscillant pour lequel le temps d'intégration est toujours supérieur à 150 ms, la dynamique approche alors 10 sur une bande de 1 GHz.

Pour améliorer le système, il existe une solution simple, utilisée depuis longtemps dans les fréquencesmètres du commerce. Elle consiste à mesurer la période et à calculer l'inverse; il faudrait bien évidemment l'adapter au cas des détecteurs, mais si par exemple on compte non seulement les impulsions entre le temps t_0 et t_1 mais aussi la durée entre la première et la dernière impulsion, la dynamique du système d'acquisition s'améliore de façon spectaculaire: pour perdre moins de 10 % sur le temps d'intégration, il suffit qu'il y ait plus de 10 impulsions entre t_0 et t_1 , ce qui donne une dynamique de 1000 pour 10 ms d'intégration.

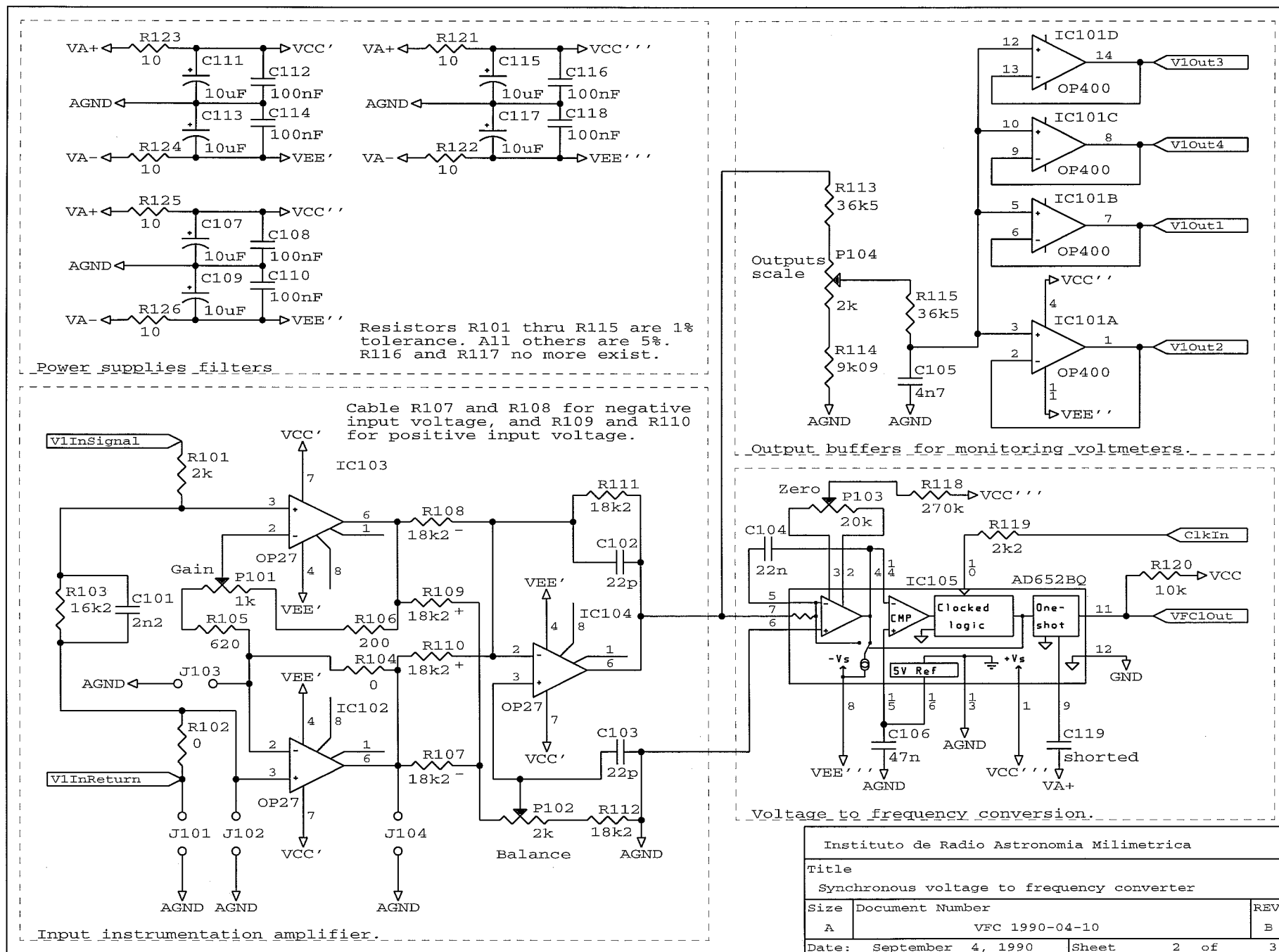
Si l'amélioration de l'interface n'a pas été encore réalisée, c'est tout d'abord parce qu'elle n'est pas urgente dans les conditions actuelles d'utilisation et qu'il fallait avant tout prouver que les limitations sont réellement dues à ce facteur. De plus, l'instrumentation du télescope est dans une phase de transition, tout le système de contrôle de l'antenne et les interfaces des batteries de filtre utilisent un standard, appelé CAMAC, qui est devenu obsolète et pour lequel il vaudrait mieux éviter de gaspiller du temps et de l'argent dans la conception de nouveaux modules. Les nouveaux corrélateurs sont le premier système au télescope à utiliser le bus VME, nettement plus fiable et moderne. Mais ils ont été installés presque 2 ans après la mise en service des nouveaux détecteurs, et c'est maintenant seulement qu'on pourrait se lancer dans la conception d'un module d'acquisition spécifique pour le continuum, puisqu'il semble bien qu'il n'existe pas de solution commerciale satisfaisante à notre problème.

3.10.3 Schéma théorique

Le schéma du circuit est montré aux pages 51 et 52. Chaque carte imprimée contient deux canaux, la deuxième page du schéma étant dupliquée.

Le circuit a été prévu pour d'autres applications et offre plus de possibilités que ce qui était strictement nécessaire aux détecteurs continuum. Il se compose de trois parties:

Figure 3.6: Schéma d'un canal de conversion tension fréquence



- Un amplificateur qui amène le signal au niveau exigé par le convertisseur tension fréquence. Cet amplificateur possède des entrées différentielles pour s'affranchir des problèmes de différence de potentiel de masse entre le détecteur et le convertisseur, il est construit autour de 3 circuits OP27 selon le schéma classique dit de l'amplificateur d'instrumentation.
- Le convertisseur tension fréquence proprement dit, avec des inverseurs à hystérésis pour les signaux logiques.
- Des tampons recopiant le signal d'entrée amplifié pour des voltmètres de contrôle et des enregistreurs papier. Ces tampons sont précédés d'un pont diviseur ajusté pour donner une tension de pleine échelle de 2 V. L'amplificateur OP400 qui assure cette fonction a été choisi pour sa faible tension d'erreur et sa faculté d'accepter sans osciller la forte charge capacitive représentée par les quelques vingt mètres de câbles menant au panneau de contrôle et presque cent mètres pour ceux qui remontent jusqu'à la cabine des récepteurs.

3.10.4 Réalisation pratique

Le tableau brossé au début de ce paragraphe sur les qualités des convertisseurs tension fréquence synchrone est un peu trop idyllique: la fréquence de sortie présente une fâcheuse tendance à se verrouiller intempestivement sur un sous-harmonique de la fréquence d'horloge, Pour éliminer ce problème, il a été nécessaire de prendre les précautions suivantes:

- l'horloge externe doit présenter des flancs suffisamment raides, qui sont améliorés par les tampons à hystérésis (Schmitt-triggers) présents sur la carte. L'idéal aurait probablement été de disposer de l'horloge sous la forme d'un signal différentiel distribué par l'intermédiaire de paires torsadées.
- Les découplages doivent être très soignés, toutes les alimentations sont découplées fonction par fonction à l'aide de cellules R-C.
- Le convertisseur reçoit l'horloge à travers une résistance dont la valeur a été déterminée expérimentalement pour diminuer les bruits introduits dans le composant sans augmenter excessivement les temps de montée de et descente.
- Enfin, le fait que le signal d'entrée soit naturellement bruité aide à éviter ces problèmes une fois que les précautions précédentes ont considérablement réduit la largeur des plages de verrouillage.

L'implantation du circuit sur une carte simple Europe (100 × 160 mm) est simple et très aérée.

Quelques détails n'apparaissent pas sur le schéma, comme l'anneau de garde qui protège l'entrée inverseuse de l'intégrateur d'éventuels courants de fuite, ou le fait que les potentiomètres ont été placés de façon à pouvoir effectuer les réglages en place, sans avoir besoin d'extraire la carte ni de la placer sur un prolongateur.

3.11 Réglages

Les détecteurs ont été conçus pour pouvoir être utilisés avec les récepteurs actuels de l'IRAM de 500 MHz de bande passante mais en prévision d'une nouvelle génération disposant de 1 GHz.

De façon à ce que les spectromètres travaillent à niveau constant et limiter la dynamique nécessaire, c'est la puissance par unité de bande de fréquence qui est spécifiée: en théorie -47 dBm/MHz , ce qui fait $20 \mu\text{W}$ dans 1 GHz de bande. Cette valeur est bien adaptée aux caractéristiques du détecteur sur du bruit (voir plus loin), donc aucun dispositif n'a été prévu pour ajuster la puissance du signal radiofréquence. Cela signifie que, pour passer d'un récepteur de 500 MHz de bande à un autre de 1 GHz, il suffit de changer le gain de l'amplificateur de la carte convertisseur d'un facteur 2.

Dans la réalité, c'est la puissance sur la charge chaude qui est proche de la valeur donnée ci-dessus, la valeur sur le ciel étant assez nettement plus faible. Il est cependant nécessaire d'avoir une marge de sécurité, et il a été décidé, surtout par tradition hérité du détecteur continuum précédent, que la puissance attendue correspondrait à la moitié de la pleine échelle du convertisseur. Ce n'est certes pas le plus favorable au point de vue du bruit de quantification mais en cas de besoin on peut toujours ajuster les atténuateurs par pas dont disposent tous les récepteurs.

La tension d'erreur du premier amplificateur est ajustée pour obtenir 0 sur les voltmètres de contrôle en l'absence de signal d'entrée. Ensuite, les convertisseurs sont réglés pour obtenir entre 50 et 100 impulsions par seconde dans les mêmes conditions; cette valeur représente moins d'un dix-millième de la pleine échelle et son effet sur la calibration est négligeable, très inférieur à celui des erreurs de linéarité.

3.12 Maintenance

Au vu des caractéristiques de dérive de l'amplificateur, il est préférable d'ajuster sa tension d'erreur régulièrement, environ 2 fois par an, et par la même occasion de retoucher le réglage de la fréquence de sortie du convertisseur en l'absence de signal. Ces opérations doivent aussi être effectuées un mois après la mise en service définitive des détecteurs puisque la dérive initiale est beaucoup plus forte d'après le fabricant, en quelque sorte l'amplificateur a besoin d'une période de rodage. Sur un petit nombre de canaux, cette maintenance représente une charge de travail minime.

Signalons que la solution d'un amplificateur composite, alliant le faible bruit du LT1028 à la stabilité à long terme d'un amplificateur à découpage a été envisagée; elle aurait éliminé ou au moins réduit considérablement le besoin de réglages réguliers, mais elle a été rejetée au vu des constantes de temps nécessaires pour ne pas dégrader le bruit, de la complexité qu'elle aurait introduite et du risque de voir apparaître la fréquence de découpage dans le signal détecté.

3.13 Mesure des caractéristiques en laboratoire

3.13.1 Niveau de bruit et autres paramètres électroniques

Le niveau de bruit en l'absence de signal radiofréquence a été mesuré à l'aide d'un analyseur de spectre audio (modèle HP 3582A). Cet appareil, qui effectue l'analyse spectrale par

transformée de Fourier d'un signal et peut afficher les résultats en densité spectrale (par exemple en $\mu\text{V}/\sqrt{\text{Hz}}$) a été une aide précieuse tout au long de la mise au point du circuit. Bien que cet instrument soit capable d'analyser le signal à des fréquences très inférieures au hertz, le temps d'intégration nécessaire pour obtenir une mesure précise dans le cas de bruit est tel que la plupart des mesures ont été effectuées entre 5 et 10 Hz, avec une résolution de 1 Hz qui donne une erreur relative de 3 % en 20 mn. Le bruit propre de l'analyseur, 25 à $35 \text{ nV} \cdot \text{Hz}^{-1/2}$, n'a rien d'exceptionnel mais est négligeable après le gain en tension de 350 à 400 du détecteur.

La valeur du bruit donnée par l'analyseur de spectre est de 1.3 à $1.4 \text{ nV}/\text{Hz}^{1/2}$ à 10 Hz, en parfait accord avec les $1.36 \text{ nV}/\text{Hz}^{1/2}$ déduits des caractéristiques typiques de l'amplificateur LT1028 dans la configuration du circuit.

Une mesure sérieuse de la fréquence de remontée du bruit rose était impossible tant que le boîtier avec le contrôle de température définitif n'était pas prêt. Une tentative de mesure du bruit sur de plus longues durées a donc été effectuée sur enregistreur papier une fois la température bien stabilisée: il en ressort que le bruit pic à pic équivalent à l'entrée de l'amplificateur est très proche de 50 nV sur une heure. Si l'on estime que la bande passante de l'enregistreur est 10 Hz (sur des mouvements de plume de faible amplitude) et que le rapport du bruit pic-pic à sa valeur efficace est 6.6, critère utilisé par les fabricants d'amplificateurs et justifié par le nombre de points indépendants de l'enregistrement (on devrait pouvoir prendre 8, mais le bruit est-il gaussien à ce point là ?), on obtient en appliquant (3.7) 46 nV de bruit pic-à-pic théorique, un accord à vrai dire étonnant. Il faut admettre que cette mesure n'est pas d'une rigueur absolue, mais c'était de loin la plus délicate de toutes à réaliser, et elle ne permet pas de mesurer les fluctuations de gain du détecteur qui limitent assurément la stabilité en présence de signal radiofréquence sur des durées aussi longues. On peut cependant considérer qu'elle garantit que la fréquence de remontée du bruit rose ne dépasse pas 3 Hz.

On peut conclure de ce qui précède que le niveau minimum de puissance qu'il faut fournir au détecteur pour qu'il contribue au bruit de façon négligeable sur 1 GHz de bande passante vaut environ $0.3 \mu\text{W}$ lorsqu'on utilise un mode de commutation rapide, et $0.8 \mu\text{W}$ lorsque la fréquence de commutation est 0.5 Hz. Il faut multiplier ces estimations par 0.7 pour un récepteur de 500 MHz de bande passante.

Les quatre détecteurs sont installés dans la même boîte et alimentés en parallèle, ce qui pourrait poser des problèmes de diaphonie, surtout si l'un des récepteurs est extrêmement instable (ce qui arrive quand on les remplit d'hélium) pendant qu'on observe une source faible avec un autre. En fait, l'isolation entre deux canaux voisins est de 66.5–67 dB ramenés à l'entrée (133–134 dB mesurés sur le signal détecté), cette valeur a été obtenue en injectant un signal monochromatique modulé à 400 Hz. Elle a toutes les chances de dépendre fortement de la qualité de l'alimentation mais est de toutes façons plus que suffisante. La valeur obtenue sur le ciel, 47 dB [17], moins bonne, est due à des couplages entre les étages amplificateurs de la boîte de distribution.

3.13.2 Linéarité en puissance

C'est le point le plus critique du détecteur pour obtenir un étalonnage correct. Il est en fait indispensable d'effectuer un tri en quadraticité des diodes et de ne retenir que les meilleures, ce qui est déplorable du point de vue de la reproductibilité du circuit. C'est de plus une opération relativement délicate, qui doit être effectuée avec le plus grand soin.

Afin de se placer dans des conditions aussi proches que possibles de l'utilisation as-

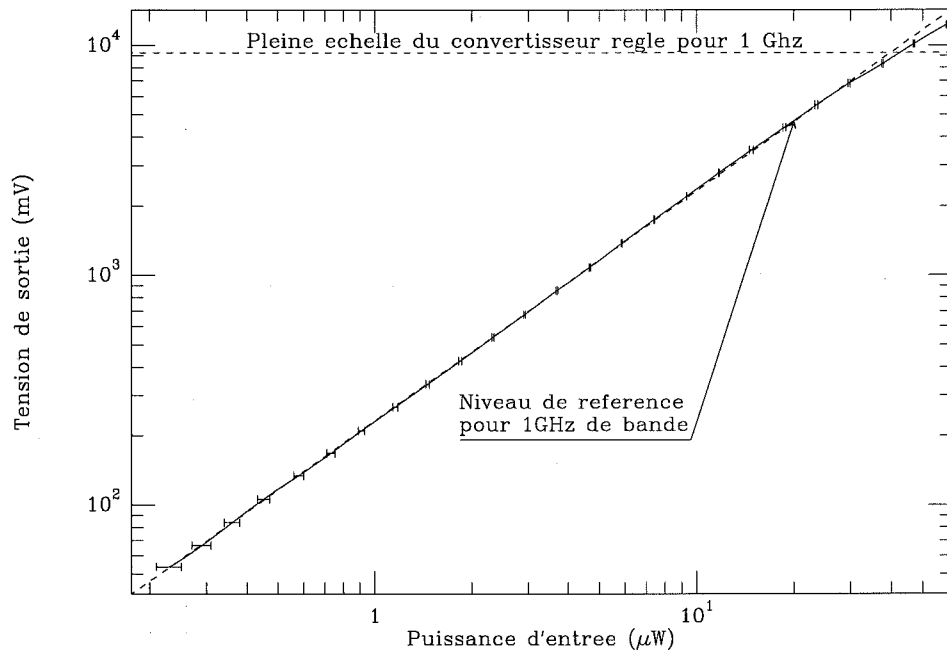


Figure 3.7: Réponse en puissance sur source de bruit.

tronomique des détecteurs, le tri est effectué à l'aide d'un générateur de bruit suivi d'un amplificateur et d'un atténuateur par pas. La mesure de l'erreur se fait par comparaison avec un wattmètre à thermocouple, seules sont gardées les diodes pour lesquelles la variation du rapport entre la tension de sortie et la puissance d'entrée entre 1 et $30 \mu\text{W}$ ne dépasse pas $\pm 1.5 \%$. Sur 45 diodes achetées en deux lots, dont 4 n'ont pas passé le cap des mises au point préliminaires, le rendement du processus de sélection s'établit à 20 %.

La réponse sur du bruit d'un détecteur est montrée en figure 3.7. La droite de régression qui est représentée est de la forme $V_s = \alpha P_i$, donc de pente 1 sur l'échelle logarithmique. Il est en général indiqué dans la littérature que les détecteurs à diode commencent à s'écarter significativement de la réponse quadratique lorsque la puissance atteint environ $10 \mu\text{W}$ et ce pour un signal monochromatique. Le graphique montre clairement que dans le cas de ce détecteur, cette puissance est proche de $30 \mu\text{W}$ sur du bruit blanc; elle atteint environ $100 \mu\text{W}$ sur un signal sinusoïdal. Ceci, joint au très faible bruit intrinsèque de l'électronique, signifie que la dynamique obtenue est très supérieure à la moyenne.

La figure 3.8 reprend les mêmes données que la précédente mais essaie de les interpréter sous la forme de l'erreur d'étalonnage due au détecteur en affichant la pente mesurée pour des écarts de puissance de 1, 3 et 6 dB. Les barres d'erreur correspondent à une unité d'affichage du wattmètre pour la mesure à 1 dB et sont plutôt optimistes quant à la qualité de ce dernier. L'un des points sort suffisamment de la courbe pour mettre en doute sa validité, cependant il suffit de changer sa valeur de 0.7 %, pour que tout rentre dans l'ordre. Ce point pourrait bien être le résultat d'une erreur de prises de notes lors des mesures. L'erreur d'étalonnage que l'on peut attendre est indiquée par la dispersion des points située à la même abscisse, qui reste inférieure à 3 % tant que la puissance maximale ne dépasse pas $25 \mu\text{W}$ pour une différence entre ciel et charge chaude comprise entre 3 et 6 dB.

L'erreur de linéarité du convertisseur tension-fréquence est quant à elle complètement

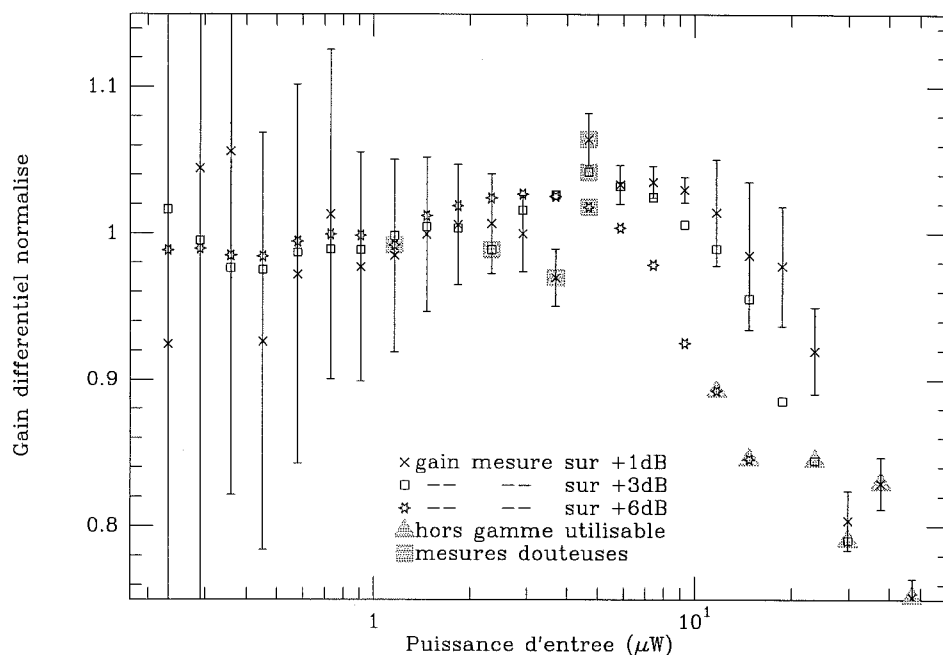


Figure 3.8: Erreur de linéarité différentielle.

négligeable. Une tentative de la mesurer à l'aide du meilleur voltmètre disponible au laboratoire (20000 points) s'est soldée par un échec: on avait plutôt l'impression de mesurer l'erreur de linéarité du voltmètre tant l'accord entre deux convertisseurs, mesuré comme le rapport des fréquences de sortie alors qu'ils travaillaient à des niveaux différents, était bon.

3.13.3 Ondulation de la réponse en fréquence

La variation de la réponse a été mesurée par comparaison du détecteur avec un wattmètre à thermocouple, la puissance au moment de la mesure était de l'ordre de $10 \mu\text{W}$ (un peu moins vers les hautes fréquences à cause du câble relativement long reliant le générateur au système de mesure). Le thermocouple était très long à se stabiliser à ce niveau de puissance après chaque changement de fréquence et on peut estimer son erreur à $\pm 0.5\%$ ($\pm 0.02 \text{ dB}$).

Le résultat de la mesure est montré à la figure 3.9. On note une légère ($0.15 \text{ dB} \approx 4\%$) perte du rendement aux environs de 1000 MHz suivie d'une remontée. Cette résonance due aux éléments parasites des composants du filtre d'entrée est suffisamment faible pour être négligée: elle représente une fraction minime de l'ondulation de la bande du récepteur après tous les étages d'amplification et de changement de fréquence. Si l'on essaie d'estimer la perte relative de temps d'intégration due à cette ondulation par rapport à une bande parfaitement plate, on obtient une valeur dérisoire, inférieure à 10^{-4} .

3.13.4 Taux d'onde stationnaire

La caractéristique de la diode (figure 3.3) montre clairement que son impédance est proche de 50Ω , et donc que, du moins tant que le niveau de puissance est suffisamment faible, le taux d'onde stationnaire reste proche de 1.

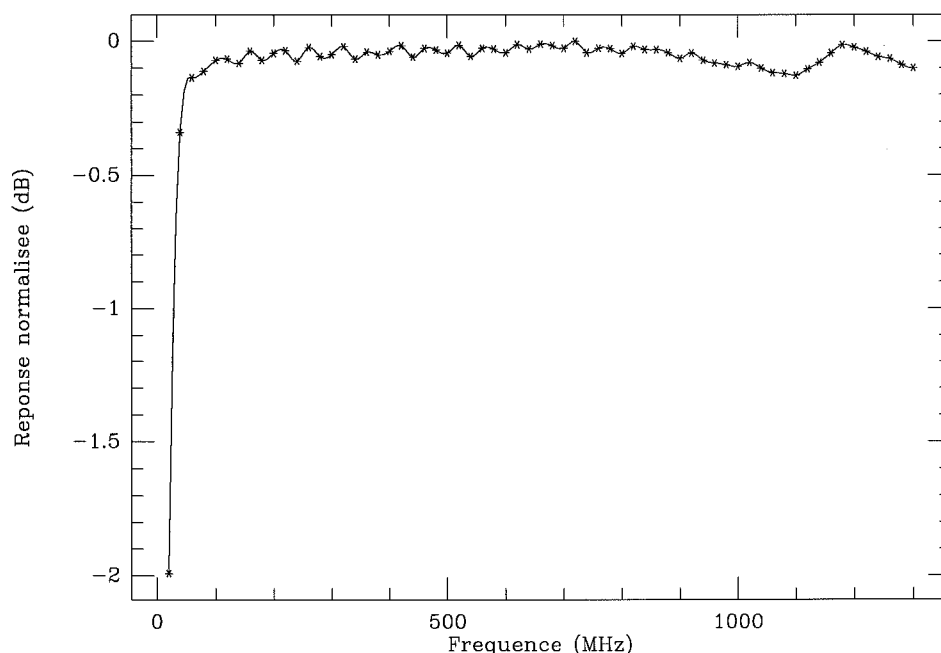


Figure 3.9: Ondulation de la réponse en fréquence du détecteur.

Ceci a été vérifié à l'aide d'un analyseur de réseau à l'Observatoire de Meudon sur un prototype du détecteur qui ne possédait pas la version finale du filtre d'entrée: le taux d'onde stationnaire restait en dessous de 1.2 entre 300 et 2400 MHz, du moins pour des puissances inférieures à $100\text{--}200\text{ }\mu\text{W}$. Il faut cependant se méfier de ce dernier chiffre, en effet, le circuit étant fortement non-linéaire, le signal réfléchi est riche en harmoniques du signal incident (essentiellement l'harmonique 2 au niveau de travail habituel) et l'analyseur utilisé étant de type vectoriel, il ne mesure que la puissance réfléchie à la fréquence de l'onde incidente et donc la sous-estime. Il n'en demeure pas moins que l'on peut obtenir facilement, sur une large bande, une très bonne adaptation d'impédance dans un système $50\text{ }\Omega$, et ce sans gaspiller 95 % ou plus de l'énergie incidente dans une résistance parallèle.

3.14 Mesures au télescope

Pour un certain nombre de paramètres les mesures effectuées en laboratoire étaient insuffisantes ou entachées d'une trop grande incertitude et il a fallu attendre l'installation définitive pour pouvoir utiliser le système d'acquisition du télescope et effectuer les mesures qui nécessitaient soit un très long temps d'intégration, soit une très grande dynamique. Par exemple, la théorie indique que le densité spectrale du signal détecté sur du bruit est -90 dBV/Hz en dessous de la tension continue, et cette valeur est suffisamment proche de la limite de dynamique de l'analyseur de spectre audio pour obliger à utiliser un couplage alternatif, interdisant toute mesure valable en dessous de 1 Hz. De plus, les expériences sur table ont, à de rares exceptions près, été effectuées sans le système de contrôle de température, ce qui rend douteuses les mesures de stabilité.

Pour toutes les figures montrées ici, les détecteurs étaient réglés pour travailler avec des

récepteurs de 500 MHz de bande passante. Donc la fréquence de sortie des convertisseurs valait 20 kHz/ μ W.

3.14.1 Variances d'Allan

Toutes les variances d'Allan présentées dans cette section correspondent à des séquences de quatre échantillons avec les poids $\{-1, 1, 1, -1\}$. Il est logique de procéder ainsi car les observations en mode continuum reposent toujours sur un grand nombre d'échantillons, dont les propriétés de filtrage des instabilités sont beaucoup plus proches de cette séquence que de la séquence $\{-1, 1\}$ et les procédures de réduction de pointage et des procédures « ONOFF » éliminent systématiquement les dérives au premier ordre. Le temps d'intégration de base est 50 ms et par conséquent la durée de cycle la plus courte vaut 200 ms.

Sur la figure 3.10, les deux courbes supérieures montrent la variance d'Allan de deux détecteurs mesurés simultanément. Le signal provient une source de bruit fournissant environ 12 μ W. Les courbes commencent à se détacher de la ligne attendue vers 5 s de temps de cycle, bien qu'elle continue à descendre jusque vers 30 s. La différence vers les très faibles temps de cycle s'explique par le fait que la source de bruit n'est pas filtrée et que la bande passante des diviseurs de puissance de la boîte de distribution s'étend bien au delà de 1.1 GHz; cependant, au delà de cette limite, la réponse en fréquence peut être très différente d'un canal à l'autre. De plus, peu après avoir effectué ces mesures, nous nous sommes rendus compte du fait que les diviseurs de puissance de la boîte de distribution pouvaient souffrir de légères instabilités de gain et donc dégrader les résultats.

Il est cependant clair que, étant donnée la similitude des détails de la courbe, les deux détecteurs sont très fortement corrélés et que la déviation du comportement idéal peut être due à une instabilité de la source de bruit.

Pour essayer d'éclaircir ce problème, la courbe inférieure présente la variance d'Allan du rapport des deux détecteurs: si les deux signaux étaient indépendants, la variance du rapport serait égale à la somme des variances individuelles. Comme c'est loin d'être le cas, il faut en conclure que les instabilités observées proviennent presque exclusivement de la source de bruit, et que les valeurs de temps de stabilité données au paragraphe précédent sont franchement pessimistes; il est malheureusement impossible de les estimer avec une précision raisonnable. Si l'on suppose que le rapport entre les courbes supérieures et inférieures, qui est compris entre 10 et 20, est causé entièrement par les défauts des détecteurs, cela signifie qu'ils ne causent pas une perte de temps d'intégration supérieure à de 5 %; pour les raisons expliquées plus haut, c'est probablement une estimation très pessimiste.

La figure 3.11 montre les variances d'Allan du premier canal et du rapport entre les deux canaux dans des conditions légèrement différente du système d'acquisition, alors que la source de bruit fournissait une puissance d'environ 2.5 μ W. Les deux courbes supérieures sont similaires et très proches de celle de la figure 3.10, le comportement du détecteur est donc satisfaisant même après avoir diminué la puissance de 7 dB.

La courbe médiane a été obtenue en laissant le système d'acquisition dans son mode normal. La courbe inférieure a été obtenue en supprimant le signal d'inhibition des compteurs. La différence observée demande évidemment quelques explications qui sortent du domaine du détecteur proprement dit et nous amènent à parler de détails de l'interface avec les modules de comptage d'impulsions.

Comme mentionné auparavant, les compteurs d'impulsions sont lus toutes les 50 ms, ces compteurs disposent d'une entrée d'inhibition qui normalement est activée pendant 60 μ s par le système de contrôle du télescope au moment de chaque interruption du microprocesseur

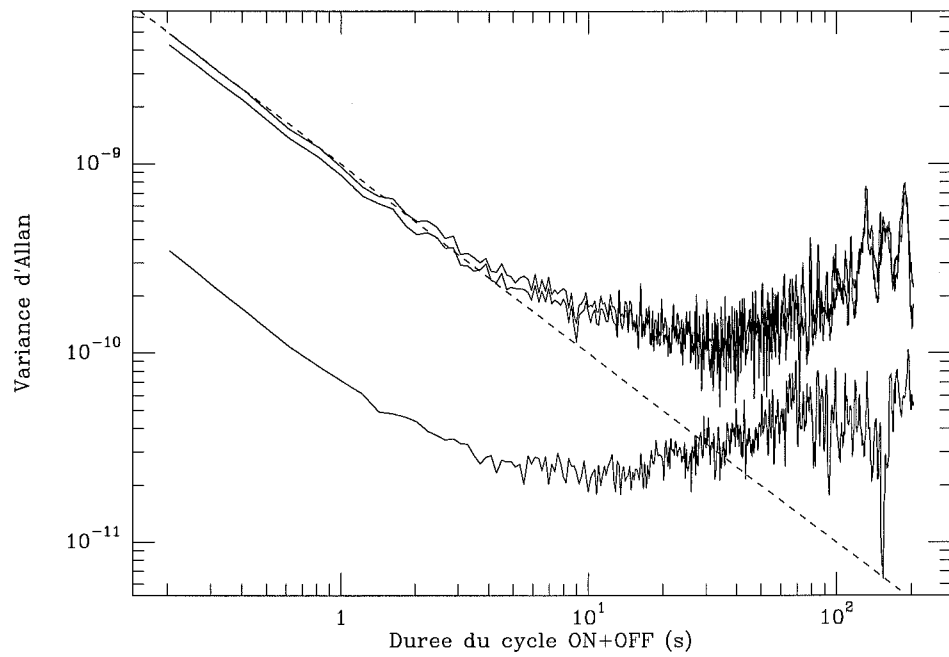


Figure 3.10: Variance d'Allan à fort niveau

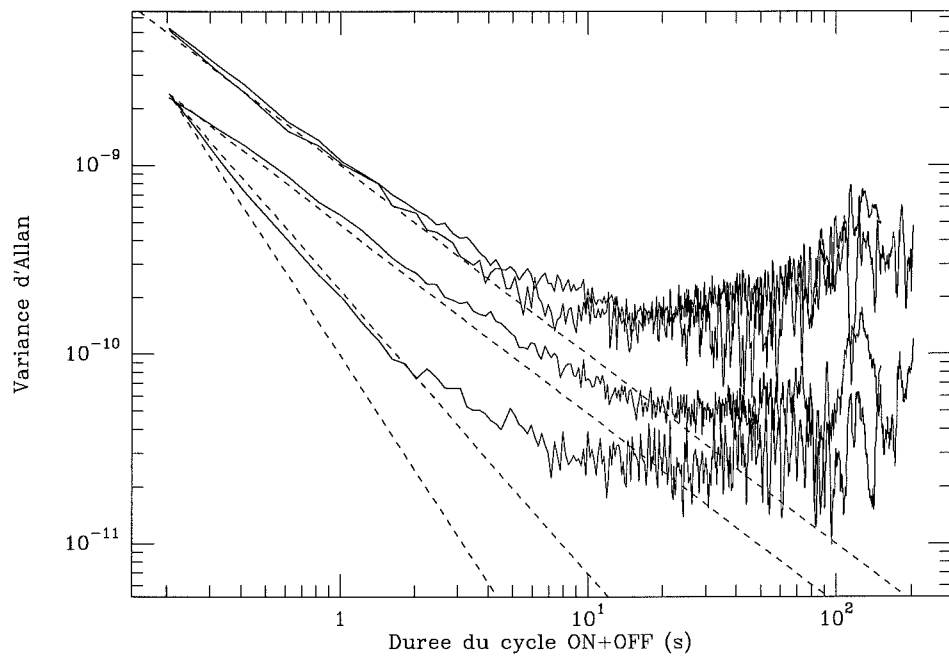


Figure 3.11: Variance d'Allan à faible niveau montrant l'influence du bruit de quantification (voir texte)

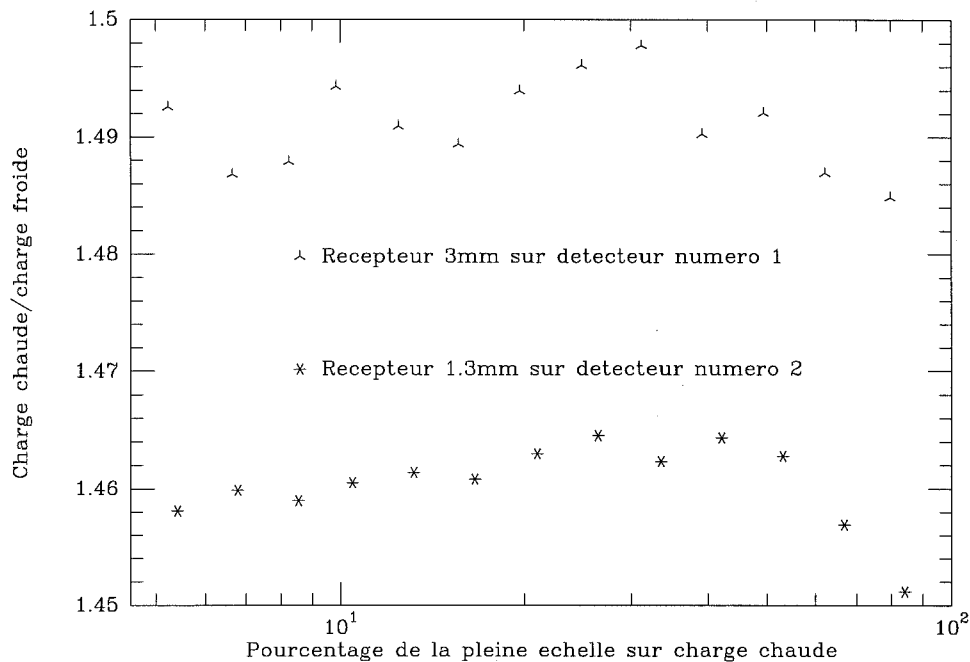


Figure 3.12: Facteur d'étalonnage en fonction de la puissance.

qui va lire les compteurs. Dans ce cas, étant donné que la constante de temps du détecteur est à peu près $15 \mu\text{s}$, on peut considérer que deux lectures successives sont indépendantes. Par contre, si on supprime ce signal d'inhibition, la durée pendant laquelle les compteurs sont aveugles lors de l'interruption de chargement et effacement devient très courte, environ 200 ns. Dans ce cas, lors du calcul de la variance d'Allan, les sommes d'échantillons adjacents possèdent un bruit de quantification égal à celui d'une seule mesure de durée égale à la somme des durées individuelles. C'est ce qui explique l'amélioration de la corrélation entre les canaux et démontre clairement que même à ce niveau, au moins 10 dB en dessous de la saturation des diodes, le détecteur est limité par le système d'acquisition. Si l'on diminuait encore le niveau ou que l'on diminuait le temps d'intégration, la pente initiale se rapprocherait de -2 , valeur qui indiquerait que l'incertitude est totalement dominée par la quantification de la mesure.

3.14.2 Linéarité en puissance

Les mesures de linéarité différentielle par comparaison avec un wattmètre étaient limitées par l'incertitude de ce dernier. Celles-ci ont été répétées au télescope en effectuant une série de douze étalonnages successifs et en diminuant la puissance de 1 dB avant chaque répétition de la procédure.

Le résultat de ces mesures pour deux récepteurs, sous la forme du rapport Y des puissances mesurées entre la charge chaude et la charge froide, est montré à la figure 3.12. L'erreur relative d'étalonnage est donnée par la variation de $Y - 1$, qui est constant à $\pm 1.5\%$ près malgré le début de saturation visible vers les fortes puissances. Cette légère compression ($\approx 0.5\% = 0.02 \text{ dB}$) apparaît un peu plus bas que ce que la comparaison avec le wattmètre laissait penser; elle pourrait être due aux étages amplificateurs de la boîte

de distribution qui précèdent le détecteur, il est cependant bien difficile de l'affirmer étant donné la précision de mesure nécessaire, qui relève presque de l'utopique. Remarquons au passage que le premier récepteur était légèrement instable au cours de cette expérience et que leur température de bruit n'était franchement pas bonne.

La conclusion que l'on peut tirer de ces mesures est que les détecteurs ont une dynamique comprise entre 30 et 100 dans les configurations utilisées en pratique, c'est à dire avec des durées de cycle n'excédant pas 2 s, et que cette dynamique est limitée par le bruit de quantification des convertisseurs tension fréquence avec l'interface actuelle.

3.15 Exemples d'observations

3.15.1 Pointage

Les figures 3.13 et 3.14 correspondent respectivement au dernier pointage effectué avec l'ancien détecteur continuum et au premier réalisé avec le nouveau. Les deux observations ont été effectuées sur la même source à quelques heures d'intervalle (le temps d'installer et de connecter la nouvelle boîte de distribution), mais à peu près à la même élévation et dans des conditions comparables. Le reste se passe de commentaires.

3.15.2 Détermination du flux d'un quasar

Un exemple de procédure « ONOFF » est donné en figure 3.15. Cette observation a été réalisée avec H. Steppe dans de bonnes conditions en Janvier 1991, avec le récepteur SIS 3mm réglé en simple bande à environ 105 GHz. Le rapport signal sur bruit est de 100 pour un flux estimé un peu supérieur à 1 Jy, ce qui donne un bruit équivalent de $170\text{--}180 \text{ mJy} \cdot \text{Hz}^{-1/2}$, plus faible d'un bon facteur 2 (en gros 5 en temps d'intégration) que les meilleurs chiffres donnés jusqu'alors.

Cette observation a été effectuée avec le miroir secondaire commutant entre la source et la position de référence avec une période de deux secondes. Cette durée relativement longue est celle qui donne les meilleurs résultats car elle rejette la modulation à 1 Hz de la puissance issue du récepteur par la pompe du circuit d'hélium et minimise quelque peu le temps perdu en mouvement du miroir, et ce bien qu'elle ne soit plutôt défavorable si l'on considère les fluctuations atmosphériques et les variations de gain du récepteur.

3.16 Mesure de la stabilité des récepteurs

Un exemple de mesure de stabilité de plusieurs récepteurs est donné à l'appendice A. On y remarquera que, malgré le très faible temps d'intégration, la remontée du bruit de quantification vers les hautes fréquences n'est visible que sur la source de bruit.

Signalons simplement que depuis ce rapport, écrit en Juin 1991, il est devenu apparent que c'est la source de bruit dont la stabilité s'est dégradée; c'est peut être dû à un problème dans l'alimentation, les diodes à bruit demandant une excellente régulation, de préférence en courant, pour être stables.

74; 2 1226+023C2 SIO V1(2-1) IRAM-30M-B11 O: 26-JUN-1990 R: 15-MAR-1991
RA: 12:26:33.249 DEC: 2:19:43.30 (1950.0) Offs: 0.0 0.0 Eq
Unknown Tau: 0.1061 Tsys: 520.3 Time: 1.6667E-02 El: 30.25
N: 134 IO: 1.000 AO: -68.15 Da: 1.025
FO: 86243.3500 Df: 500.0 Pos. Ang. : 90.00 Ho
B ef: 0.9300 F ef: 0.9300 G im: 0.2500
6436

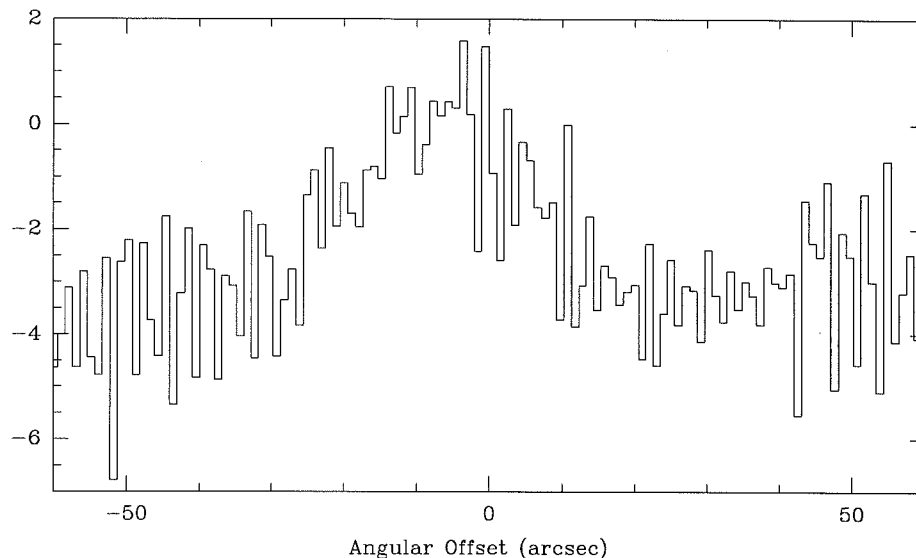


Figure 3.13: Dernier pointage avec l'ancien détecteur ...

139; 4 1226+023C2 SIO V1(2-1) IRAM-30M-B11 O: 26-JUN-1990 R: 15-MAR-1991
RA: 12:26:33.249 DEC: 2:19:43.30 (1950.0) Offs: 0.0 0.0 Eq
Unknown Tau: 8.4000E-02 Tsys: 472.4 Time: 1.6667E-02 El: 38.93
N: 99 IO: 1.000 AO: -66.58 Da: 1.359
FO: 86243.3500 Df: 500.0 Pos. Ang. : 90.00 Ho
B ef: 0.9000 F ef: 0.9000 G im: 0.2500
6460

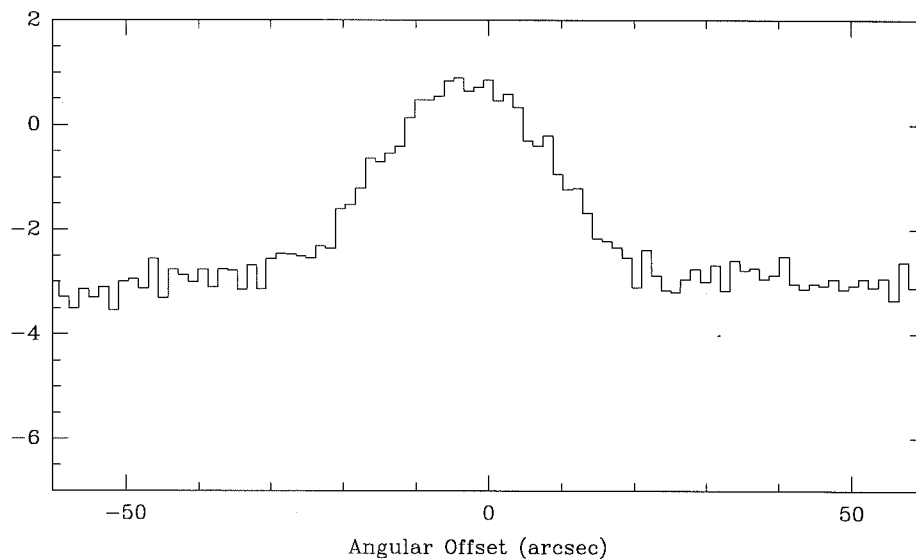


Figure 3.14: et premier pointage avec le nouveau.

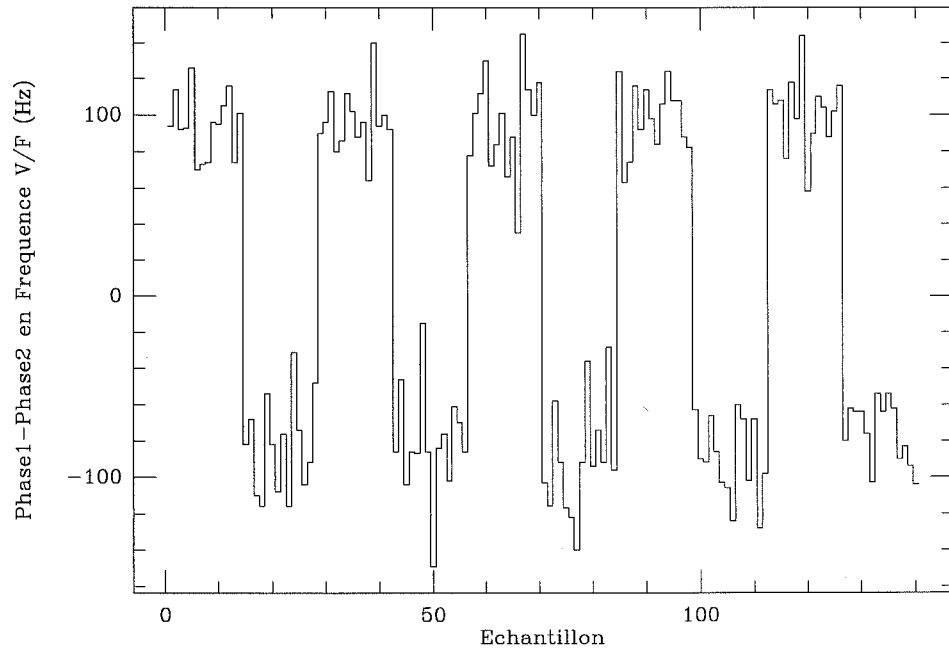


Figure 3.15: Exemple de procédure « ONOFF ».

3.17 Holographie

Il s'agit là d'une procédure qui n'est jamais réalisée par l'observateur qui vient au télescope et qui pour cela n'a pas trouvé sa place au début de ce chapitre.

Le but de l'holographie est de déterminer les erreurs de positionnement des panneaux qui constituent la surface de l'antenne. La procédure utilisée au 30m à l'heure actuelle consiste à installer un émetteur sur le Pico Veleta et à cartographier le lobe de l'antenne pour différentes positions du miroir secondaire.

Une session d'holographie a eu lieu au mois d'Août 1990 [18], avec un nouvel émetteur plus puissant et plus stable. Pour des raisons de dynamique, un détecteur spécial, mesurant non la valeur quadratique moyenne mais la valeur absolue moyenne du champ électrique, avait aussi été réalisé pour l'occasion. Les données ont été acquises en parallèle sur un détecteur continuum et le détecteur linéaire, numérisé lui aussi par l'intermédiaire d'un des convertisseurs tension-fréquence de la boîte de distribution.

La comparaison des cartes obtenues avec les deux systèmes a montré que les détecteurs continuum offrent une dynamique suffisante pour l'holographie, même si la valeur annoncée de 70 dB par l'auteur du rapport semble pêcher par optimisme, puisque pour 300 ms de temps d'intégration, la limite imposée par le bruit de quantification du convertisseur est environ 58 dB.

3.18 Remarques

On pourrait penser que le critère des 10% systématiquement choisi pour le facteur de dégradation du temps d'intégration est bien sévère. Il faut cependant tenir compte du fait qu'il ya

plusieurs contributions à l'excès de bruit, qui s'additionnent linéairement lorsqu'elles sont faibles: le bruit du détecteur, les intermodulations d'ordre supérieur à 2, le bruit de quantification du convertisseur et l'ondulation de la bande du récepteur. Cette somme devient vite inacceptable pour l'astronome quand on sait le temps qu'il passe au pointage s'il n'a pas la chance d'avoir une source forte dans la région qu'il veut observer. Ce critère correspond aussi à peu près à la précision des mesures que l'on peut réaliser lors de la mise au point du détecteur.

La sensibilité limite donnée par le fabricant pour les diodes utilisées vaut -48 dBm alors que si on la recalcule dans les mêmes conditions de bande passante dans le détecteur réalisé ici, on obtient à peu près -50 dBm, ce qui correspond à un niveau de bruit 4 dB plus faible dans la partie vidéo. Cette écart est très supérieur aux erreurs de mesure et aux dispersions sur les composants, et obtenir un circuit donnant de façon reproductible des caractéristiques nettement meilleures que celles, rarement pessimistes, annoncées par le fabricant est assez rare pour être signalé. C'est d'autant plus vrai que ce chiffre est obtenu par à partir de mesures effectuée à très basse fréquence, où le bruit rose peut jouer un rôle important et où les composants micro-ondes ne brillent en général pas et ne sont pas spécifiés.

Il a déjà été dit que l'amplificateur opérationnel à très faible bruit est un modèle récent, qui n'existait pas lors de la construction du précédent détecteur continuum; c'est aussi vrai du convertisseur tension-fréquence. Quant à la diode, s'il existait alors des modèles comparables dans le catalogue du même fabricant, la capacité parasite était sensiblement plus forte et la dispersion sur les caractéristiques était beaucoup plus importante: le courant de pic, qui est directement relié à l'impédance, était donné à $\pm 30\%$ au lieu de $\pm 5\%$. Autrement dit, aucun des trois composants les plus critiques du détecteur n'était vraiment disponible lorsque le précédent fut construit.

On pourrait aussi se demander pourquoi on n'utilise pas directement la somme des canaux d'un spectrographe à large bande comme détecteur continuum. C'est assez simple, personne n'a pour le moment réussi à obtenir la stabilité nécessaire; pour une batterie de filtres, la présence d'un seul mauvais canal suffirait à tout dégrader, pour un spectrographe acousto-optique ou un corrélateur, d'autres problèmes techniques se posent. Enfin, il est à peu près nécessaire d'avoir un détecteur en permanence qui mesure la puissance totale dans la bande de chaque récepteur, c'est très utile lors des réglages; pourquoi alors ne pas en faire un instrument de mesure avec les meilleures caractéristiques possible quand le circuit final est, tout compte fait, extrêmement simple ?

Enfin il faut dire que cette étude n'a été entreprise qu'après avoir essayé sans succès de trouver un détecteur commercial répondant au cahier des charges, assez exigeant, de l'intégration sur une bande de 1 GHz ou plus. Elle a été menée dans le cadre d'une refonte du système de distribution de fréquence intermédiaire qui était devenu insuffisant. L'ancien système était prévu pour 2 récepteurs et 4 spectromètres, la nouvelle boîte de distribution permet d'utiliser jusqu'à 4 récepteurs et 9 spectromètres ou parties de spectromètres puisque les batteries de filtre et les corrélateurs peuvent être divisées en sous bandes connectées à différents récepteur. L'installation de cette nouvelle boîte de distribution [17] a augmenté considérablement la flexibilité du télescope.

3.19 Conclusion et perspectives

L'installation de ces détecteurs au télescope a représenté un progrès important, mais il faut signaler que les détecteurs précédents s'étaient dégradés après plusieurs années d'utilisation

et que l'on ne peut a priori exclure un phénomène similaire pour les nouveaux. Cependant, leur marge de sécurité dans les conditions normales d'utilisation permet de tolérer une légère dégradation sans conséquences visibles pour l'utilisateur du télescope. En bref, le télescope dispose maintenant de détecteurs continuum adaptés à ses besoins. Le niveau de bruit obtenu est à limite de ce qui est physiquement réalisable, mais le circuit lui-même est resté très simple.

En ce qui concerne l'avenir, la disponibilité probable à moyen terme d'amplificateurs opérant vers 100 GHz avec une bande passante de 10 à 15 GHz conduirait en théorie à une amélioration notable de la sensibilité en mode continuum. Après conversion vers une fréquence intermédiaire de 2 à 18 GHz (bande la plus large pour laquelle on trouve des amplificateurs commerciaux à des prix raisonnables), des détecteurs basés sur les mêmes composants devraient offrir des caractéristiques satisfaisantes; il faudra cependant veiller à améliorer le système d'acquisition pour diminuer le bruit de quantification.

Chapitre 4

Spectrographe acousto-optique

4.1 Introduction

On a déjà signalé qu'il s'agit, avec les batteries de filtres et les corrélateurs, de l'un des trois types principaux de spectrographes utilisés en radioastronomie.

Le principe physique de base est connu depuis Brewster qui avait dès 1815 constaté qu'un matériau diélectrique soumis à une contrainte voit son indice de réfraction varier, plus précisément qu'il devient biréfringent.

Le phénomène de diffraction d'un faisceau lumineux par une onde acoustique avait été prévu par Brillouin en 1921, et vérifié expérimentalement presque dix ans plus tard.

Ce n'est cependant que dans les années 60, que les premières expériences pratiques d'application à la spectrographie ont pu être réalisées grâce à la disponibilité des lasers et, dans une moindre mesure, aux progrès en matière de fabrication de monocristaux de grande dimensions puisque certaines expériences furent réalisées avec de l'eau comme milieu de propagation.

4.2 L'interaction acousto-optique

Il ya trois façons de voir le phénomène acousto-optique qui sont passées rapidement en revue dans les paragraphes suivants. Ce sont l'optique classique, la mécanique corpusculaire et l'électrodynamique dans la matière.

4.2.1 Point de vue de l'optique classique

Considérons un milieu diélectrique dans lequel se propage une onde acoustique plane et monochromatique de pulsation Ω se propageant à la vitesse du son c_s . Supposons que cette onde provoque une modulation de la densité du milieu et que son indice dépende linéairement de cette densité. On voit que l'on réalise un réseau de pas $\Lambda = 2\pi c_s/\Omega$. Si maintenant on éclaire ce réseau par une onde monochromatique plane de longueur d'onde λ sous l'angle d'incidence θ comme illustré sur la figure 4.2.1, on voit que le front d'onde va être progressivement déformé lors de la traversée du milieu. Dans la limite d'un milieu de propagation infini, il est clair qu'une condition nécessaire pour qu'il ait propagation d'énergie dans la direction φ est qu'elle vérifie:

$$\Lambda(\sin \theta + \sin \varphi) = p\lambda$$

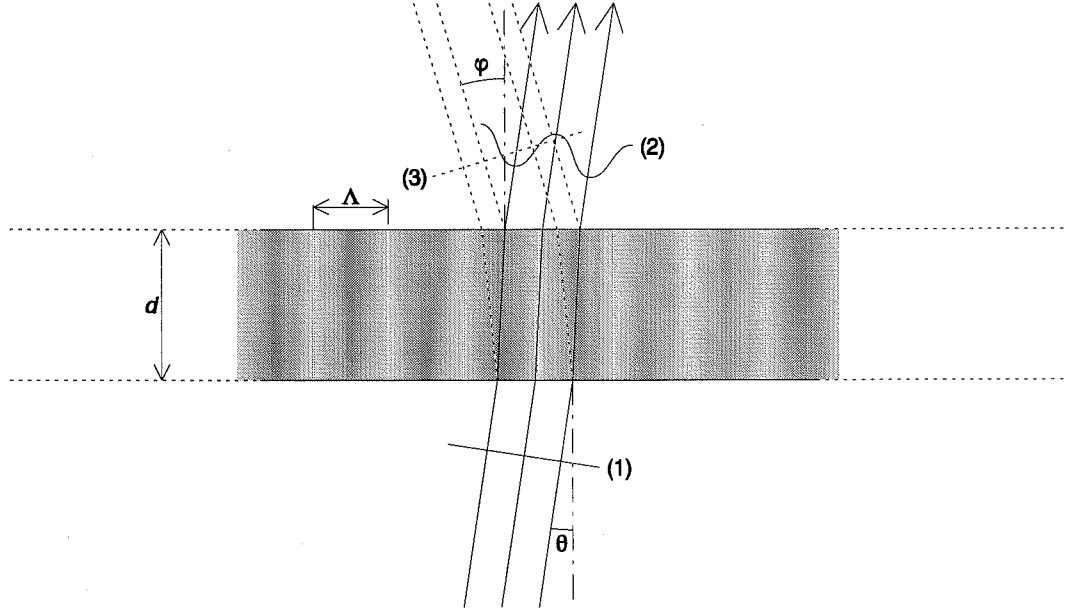


Figure 4.1: Modèle classique de l'interaction acousto-optique.

ou p est un entier arbitraire, positif ou négatif qui est l'ordre de diffraction correspondant.

On peut alors, par diverses méthodes [34, chapitre 12], résoudre sous des hypothèses raisonnables les équations de propagation dans ce système. On peut même exprimer, mais les calculs sont vraiment lourds, l'intensité des premiers ordres diffractés en fonction de quelques paramètres qui sont Λ , λ , θ , la largeur d du faisceau acoustique et son intensité.

L'avantage de ce point de vue est le fait qu'on reste assez proche de la théorie de la diffraction par un réseau. L'inconvénient est que l'on traite les deux ondes comme des champs scalaires, ignorant de ce fait tous les problèmes de polarisation dont on verra par la suite qu'ils sont importants. On ne voit pas non plus pourquoi une onde de cisaillement, à laquelle n'est associée aucune variation de densité, peut aussi produire le même phénomène. Ce point de vue, qui est celui adopté dans le calcul de Brillouin, ne rend correctement compte du phénomène que dans le cas d'un milieu liquide.

4.2.2 Point de vue corpusculaire

On peut aussi considérer l'interaction acousto-optique comme une collision entre un phonon et un photon, caractérisés par leurs vecteurs d'onde et pulsations notés K et Ω pour le premier, k et ω pour le second. Après la collision, qui résulte soit en l'annihilation du phonon, soit en l'émission d'un deuxième phonon identique au premier, le photon est défléchi et ses nouveaux paramètres seront notés k' et ω' .

Les conditions qui permettent l'interaction sont celles dans lesquelles on a conservation de l'énergie et de l'impulsion; la première condition s'écrit simplement:

$$\omega' = \omega \pm \Omega \quad \text{qui devient} \quad \omega' \simeq \omega$$

car en pratique $\omega \gg \Omega$, de 5 ordres de grandeur au moins. Cette approximation implique aussi l'égalité des modules des vecteurs d'onde:

$$|k'| = n\omega'/c \simeq n\omega/c = |k|$$

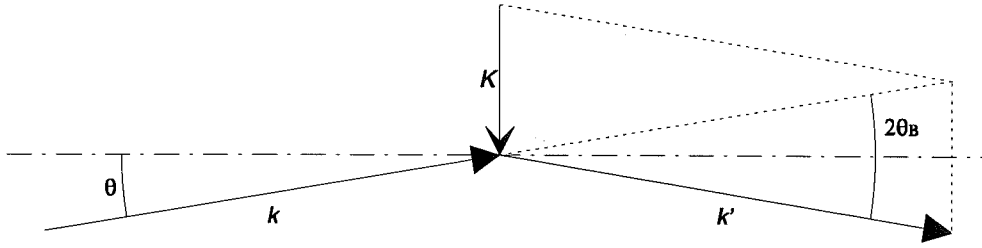


Figure 4.2: Modèle corpusculaire de l'interaction acousto-optique.

où n est l'indice du milieu supposé identique pour k et k' . C'est le cas représenté sur la figure 4.2.2. La condition de conservation de la quantité de mouvement devient alors une condition sur l'angle d'incidence θ du photon par rapport au plan d'onde des phonons qui doit être égal à l'angle de Bragg θ_B défini par:

$$\sin \theta_B = \frac{K}{2k}$$

cet angle d'incidence correspond au maximum de l'efficacité de l'interaction et privilégie l'un des faisceaux diffractés au détriment de tous les autres. On peut aussi effectuer le même calcul dans le cas d'un matériau biréfringent, pour lequel le photon incident et le photon défléchi voient deux indices différents n et n' , et montrer que pour une valeur donnée θ , il y a dans certains cas deux valeurs de K qui satisfont les conditions de conservation [35, pages 8–9].

L'avantage principal de cette approche est évidemment sa simplicité, tant en ce qui concerne les calculs que l'image physique du phénomène. On peut aussi concevoir que la section efficace de collision dépende des polarisations des deux particules. Obtenir les conditions de Bragg implique une largeur du faisceau acoustique d suffisante pour que la direction du phonon soit définie à mieux que $2\theta_B$ près, si ce n'est pas le cas, on se trouve dans les conditions dites de Raman-Nath pour lesquelles un nombre important d'ordres de diffraction est présent. Il est par contre difficile dans un modèle aussi simple de comprendre pourquoi la polarisation du photon tourne de 90° dans certains cas.

4.2.3 Point de vue électrodynamique dans la matière

C'est le point de vue le plus sophistiqué, et le seul qui permette d'effectuer les calculs valides. Comme les matériaux utilisés en pratique sont des cristaux, il est nécessaire de commencer par rappeler les propriétés de base de la propagation des ondes acoustiques et électromagnétiques dans les matériaux anisotropes.

Les contraintes et les déformations dans un milieu continu sont décrites par des tenseurs symétriques d'ordre 2 que nous noterons σ_{ij} pour les premières et u_{ij} pour les secondes, que l'on peut considérer comme la partie symétrique du gradient du vecteur des déplacements dans le matériau. Pour de petites déformations, une relation linéaire, dite loi de Hooke, existe entre ces deux tenseurs qui est de la forme:

$$\sigma_{ij} = \lambda_{ijkl} u_{kl}$$

le tenseur λ_{ijkl} , appelé tenseur des modules d'élasticité possède les propriétés de symétrie suivantes:

$$\lambda_{ijkl} = \lambda_{jikl} = \lambda_{ijlk} = \lambda_{klij}$$

ce qui ramène dans le cas le plus général le nombre de ses composantes indépendantes à 21. Cette dernière valeur n'est atteinte que dans le cas du système triclinique. Dans tous les autres cas, l'application des transformations du groupe ponctuel auquel appartient le matériau considéré permet de réduire ce nombre.

Par exemple, dans le cas du système cubique, les seules composantes non nulles sont:

$$\lambda_{xxxx} = \lambda_{yyyy} = \lambda_{zzzz}$$

$$\lambda_{yyzz} = \lambda_{zzxx} = \lambda_{xxyy}$$

$$\lambda_{yzyz} = \lambda_{zxzx} = \lambda_{xyxy}$$

soit 3 composantes indépendantes, et dans le cas des matériaux isotropes, on a en plus $2\lambda_{xyxy} = \lambda_{xxxx} - \lambda_{xxyy}$.

Tous les tenseurs d'ordre supérieur à 2 utilisés en pratique sont symétriques sur des paires d'indice, que l'on remplace pour alléger les notations par un seul indice variant de 1 à 6 avec la convention:

$$\begin{aligned} xx &\mapsto 1, & yy &\mapsto 2, & zz &\mapsto 3, \\ yz &\mapsto 4, & zx &\mapsto 5, & xy &\mapsto 6 \end{aligned}$$

qui présente l'avantage de limiter à 2 l'ordre de tous les tenseurs utilisés, ce qui permet de les représenter simplement comme des matrices. Cependant, pour que les expressions gardent la même forme, il faut faire une exception à cette règle et multiplier les composantes non diagonales du tenseur des déformations par 2: $u_4 = 2u_{yz}, \dots$

Lorsqu'on résout l'équation du mouvement dans le cas de la propagation d'une onde plane monochromatique de vecteur d'onde K_i dans ce milieu, on est amené à diagonaliser le tenseur dit tenseur de Christoffel Γ_{ij} défini par:

$$\Gamma_{ij} = \lambda_{ijkl} K_k K_l$$

ce tenseur est évidemment symétrique et possède, sauf cas de dégénérescence, trois valeurs propres distinctes qui sont les carrés des pulsations Ω correspondant à trois polarisations (direction du mouvement des particules) possibles définies par les vecteurs propres correspondant et mutuellement orthogonaux. Les vitesses de phase $|V_\phi| = \Omega/|K|$ peut alors être calculée; c'est la racine d'une combinaison linéaire des λ_{ijkl} divisée par la densité ρ du cristal. Quant à la vitesse de groupe de composantes $V_{gi} = d\Omega/dK_i$, elle n'est parallèle à V_ϕ que dans des directions privilégiées. L'angle entre le plan d'onde et la direction de propagation de l'énergie peut dépasser 60° et varier très rapidement en fonction de la direction de K [36, figure 3.22]. Le choix de la direction du faisceau acoustique permet de contrôler sa dispersion, bien que de façon assez limitée en pratique.

Quant aux propriétés optiques des cristaux, elles sont données par le tenseur diélectrique ϵ_{ij} . Ce tenseur est symétrique; on peut donc toujours se ramener à une base dans laquelle il est diagonal. On définit aussi le tenseur B_{ij} des perméabilités diélectriques comme l'inverse de ϵ_{ij} : $B_{ij}\epsilon_{jk} = \delta_{jk}$. La surface de l'ellipsoïde des indices est alors simplement la solution de l'équation $B_{ij}x_i x_j = 1$.

Une déformation perturbe B_{ij} , ce couplage peut-être décrit par le tenseur photoélastique du quatrième ordre p_{ijkl} et la relation:

$$B_{ij} = B_{ij}^0 + p_{ijkl} u_{kl}$$

d'où il ressort immédiatement que: $p_{ijkl} = p_{jikl} = p_{ijlk}$, mais il n'y aucune relation de symétrie entre les deux paires d'indices, ce qui donne jusqu'à 36 composantes indépendantes.

Connaissant maintenant les caractéristiques de l'onde acoustique et de la perturbation associée des perméabilités, on peut résoudre les équations de propagation dans le milieu. Le calcul dans les conditions de Bragg [36, paragraphe 7.8] montre que l'intensité du faisceau diffracté est proportionnelle au produit des intensités des faisceaux acoustique et optique incident et fait intervenir un facteur de mérite M donné par:

$$M = \frac{n^6 p^2}{\rho V_\phi^3}$$

où p est l'élément ou la combinaison linéaire des p_{ijkl} correspondant à la géométrie de l'interaction. Ce facteur montre clairement un certain nombre de paramètres importants dans le choix du matériau, mais en oublie d'autres comme l'atténuation de l'onde acoustique.

Parmi les composantes des p_{ijkl} , il faut distinguer les termes qui produisent des éléments non diagonaux dans B_{ij} . Ces termes modifient les directions des axes principaux de ce tenseur, et ont pour conséquence de faire tourner la polarisation du faisceau diffracté de 90° par rapport au faisceau incident. Un polariseur permet alors d'éliminer le faisceau incident et simplifie grandement les problèmes de lumière parasite. Pour les structures cristallines relativement simples possédant un haut degré de symétrie comme la cubique ou même l'orthorhombique, il n'y aucun couplage entre les termes diagonaux d'un des tenseurs et les non diagonaux de l'autre, et une interaction avec changement de polarisation correspond à une onde de cisaillement; pour des symétries moins contraignantes, ce lien disparaît.

D'autres phénomènes courants dans les cristaux peuvent intervenir pour compliquer un peu plus le problème. Le premier d'entre eux est la piézoélectricité, que l'on peut décrire par un tenseur du troisième ordre e_{ijk} qui couple le tenseur des déformations u_{ij} au déplacement électrique:

$$D_i = \epsilon_{ij} E_j + e_{ijk} u_{jk}$$

où E_i sont les composantes du champ électrique dans le milieu. Ce phénomène modifie les composantes du tenseur de Christoffel dans un certain nombre de cas pratiques importants et est loin d'être négligeable [36, pages 136–166]. Le deuxième de ces phénomènes est représenté par le tenseur électrooptique r_{ijk} , qui couple le champ électrique à une perturbation du tenseur des perméabilités:

$$B_{ij} = B_{ij}^0 + r_{ijk} E_k$$

Le champ électrique à utiliser dans cette expression inclut celui créé par la piézoélectricité du cristal lors de la propagation d'une onde acoustique, quelquefois plus importantes que celles dûes au tenseur photoélastique [36, pages 366–368]; il exclut par contre la composante dûe au faisceau optique, de trop haute fréquence pour avoir une influence. Par contre, et comme tous ces phénomènes se couplent entre eux, il faut distinguer par exemple le tenseur des permittivités diélectriques donné à contrainte constante ϵ_{ij}^σ de celui donné à déformation constante ϵ_{ij}^u .

Le principal avantage de cette approche est qu'elle décrit de façon complète les mécanismes mis en jeu. On se ramène toujours à des systèmes d'équations linéaires qui se prêtent bien à des calculs sur ordinateur, en fait indispensables si l'on veut étudier un cristal sous une orientation arbitraire. Au rang des inconvénients, il faut admettre que les calculs sont lourds parce qu'ils font intervenir des tenseurs comportant un nombre important de coefficients indépendants, et qu'on perd rapidement la vision physique des phénomènes.

4.3 Application à la spectroscopie

Dans un spectrographe acousto-optique, on utilise le fait que l'intensité du faisceau diffracté dans une direction donnée est proportionnelle à la puissance acoustique à la fréquence correspondante, du moins tant qu'on ne diffracte qu'une petite proportion de la lumière incidente. Pour satisfaire les conditions de Bragg sur une large bande, il faut que le faisceau acoustique soit suffisamment divergent dans le plan de l'onde lumineuse [36, page 297]. Il suffit après cela avec une lentille, appelée lentille de Fourier, de transformer l'angle de diffraction (image à l'infini) en position sur un photodétecteur.

Il faut bien sûr aussi transformer le signal électrique en signal acoustique, le transducteur piézoélectrique qui réalise cette fonction, situé à l'une des extrémités du cristal, est l'élément le plus complexe de la cellule de Bragg. En effet, outre le cristal piézoélectrique, qui doit être conçu pour exciter le mode de propagation désiré et minimiser les autres, il faut déposer les électrodes de masse et de signal nécessaires au fonctionnement du transducteur. Le modèle [36, figure 5.8], valide sur le plan théorique, ne tient pas compte de problèmes pratiques comme le fait qu'il faut parfois ajouter des couches intermédiaires de collage si le métal de l'électrode n'adhère pas au cristal. Dans certains cas, l'électrode de masse est aussi utilisée comme adaptation d'impédance mécanique [36, pages 237–242]. La faible épaisseur du transducteur à haute fréquence, quelques microns, s'obtient par collage d'une couche épaisse réduite progressivement par un polissage long et délicat. Dans ces conditions, il ne faut pas s'étonner si la fabrication du transducteur représente une fraction importante du coût de la cellule.

Pour éviter les ondes acoustiques stationnaires dans le déflecteur, l'extrémité du cristal opposée au transducteur est taillée en biais avant de recevoir un matériau absorbant [36, figure 5.9c].

On appelle ouverture du cristal la distance dans la direction de propagation de l'onde acoustique sur laquelle s'effectue l'interaction. Elle est plus faible que la longueur du cristal pour des raisons géométriques évidentes.

L'inverse du temps de propagation à travers l'ouverture définit en première approximation la résolution en fréquence que l'on peut obtenir. Quant au nombre de points que l'on peut résoudre, les fabricants le définissent par convention comme égal à la bande passante multipliée par le temps de propagation. L'atténuation acoustique, proportionnelle au carré de la fréquence dans un cristal avec peu de défauts, limite la longueur utilisable du cristal et donc la résolution que l'on peut obtenir.

Il existe d'autres critères classiques de résolution, par exemple le critère de Rayleigh [34, page 333], mais il n'est pas franchement utilisable quand on essaie d'obtenir d'obtenir une tache focale sans maximaux secondaires. Il n'est pas non plus très réaliste quand on le pas d'échantillonnage n'est pas négligeable par rapport à la résolution et quand on sait que les spectres astronomiques ont souvent un rapport signal sur bruit inférieur à 10.

Pour le radiotélescope de 30 mètres, la résolution lors de la conception de ce spectrographe était un critère important mais non fondamental, puisque l'on dispose en de spectrographes à haute résolution (filtres et corrélateurs) que l'on peut connecter en parallèle si nécessaire pour étudier plus en détail certaines parties du spectre. Il en va autrement pour les missions spatiales où les critères d'encombrement, de poids et de consommation électrique font que l'on cherche à fabriquer le spectromètre universel, qui couvre une large bande et donne en même temps une haute résolution.

Plus précisément, et en orientant l'axe des x selon la direction de propagation du faisceau acoustique et plaçant son origine à l'extrémité du cristal où est placé le transducteur, la forme

du filtre équivalent est le carré du module de la transformée de Fourier du produit du profil d'éclairement en champ électrique de la cellule par l'amplitude du faisceau acoustique; cette dernière est proportionnelle à la racine de l'intensité acoustique I_a , exponentielle décroissante de longueur caractéristique l , $I_a(x) = I_a(0)e^{-x/l}$.

Avec un profil d'éclairement gaussien centré en x_0 de largeur w à $1/e$ en champ donc $E(x) = E(x_0) \exp(-(x - x_0)^2/w^2)$. Il est très facile de vérifier que le produit de ces deux fonctions est aussi une gaussienne de largeur w , seul son maximum s'est déplacé vers l'origine. Cette propriété de l'éclairement gaussien a une conséquence importante, elle signifie que la forme du filtre sera, en première approximation, indépendante de l'atténuation acoustique et permet d'utiliser des bandes de fréquence relatives larges malgré la variation de l'atténuation dans la bande. Sa transformée de Fourier est aussi une gaussienne, fonction à décroissance rapide qui donnera une bonne sélectivité.

Certes, la troncature due à la longueur finie du cristal ne permettra pas d'obtenir un filtre gaussien parfait, mais permettra de satisfaire un objectif important, celui d'obtenir des filtres de forme similaire sur toute la bande.

4.3.1 Matériaux utilisés

Les paramètres importants du matériau donnés par le facteur de mérite de l'interaction sont tout d'abord l'indice de réfraction et les composantes du tenseur photoélastique. Pour analyser le terme au dénominateur, il faut se rappeler que $V_\phi^2 \propto \rho^{-1}$ et tenir compte du fait qu'à une vitesse lente est en général associée une plus grande atténuation acoustique. Enfin il faut que le matériau soit transparent à la longueur d'onde du faisceau optique. Toutes ces contraintes limitent sérieusement l'éventail de choix des matériaux; les plus importants par ordre croissant de fréquence acoustique sont:

- La paratellurite TeO_2 , utilisée surtout car elle possède un mode acoustique extrêmement lent qui permet d'atteindre de très hautes résolutions, mais limité à des fréquences de l'ordre de 100 MHz.
- Le molybdate de plomb PbMoO_4 , possédant un bon compromis entre l'efficacité et l'atténuation pour des bandes de l'ordre de 200–300 MHz.
- Le phosphore de gallium GaP , qui exhibe le meilleur facteur de mérite grâce à son indice très élevé (3.3), même s'il n'est pas parfaitement transparent. C'est un semiconducteur de classe III-V comme l'arséniure de gallium, qui cristallise dans un système cubique et donc relativement simple.
- Enfin le niobate de lithium LiNbO_3 , qui est un matériau fort complexe, utilisé aussi bien pour les cellules de Bragg que pour les transducteurs. La description de ses propriétés mécaniques, optiques, électrooptiques, piézoélectriques et photoélastiques requiert 24 coefficients indépendants dans les tenseurs associés.

Il est peu probable que cette liste s'allonge rapidement malgré les efforts entrepris pour trouver de nouveaux matériaux possédant de meilleures caractéristiques acoustiques et optiques.

4.4 Le spectrographe installé à l'IRAM

4.4.1 Architecture générale

Elle est présentée à la figure 4.3; les divers éléments en sont décrits ci-dessous avec plus de précision. L'ensemble se présente sous la forme de deux racks, le premier contient l'optique et le processeur de fréquence intermédiaire, le deuxième les alimentations, l'électronique d'acquisition et le micro-ordinateur qui assure l'interface avec le reste du télescope.

4.4.2 La cellule de Bragg

C'est par elle qu'il faut commencer car les choix du matériau et de la fréquence de travail influencent à peu près tous les aspects de la construction du spectrographe.

Elle est réalisée en phosphure de gallium GaP et utilise un mode de propagation longitudinal selon la direction $\langle 1, 1, 0 \rangle$ du cristal. Son ouverture vaut 6.32 mm, ce qui, avec une vitesse acoustique proche de 6300 m/s [36, page 198] (mais si on la recalcule à partir des valeurs données en appendice, on obtient 6450 m/s), donne un temps d'interaction très proche d'une microseconde. Sa fréquence centrale est 1 GHz avec une bande passante à -3 dB de 510–520 MHz.

Bien évidemment, la cellule et tous les composants optiques possèdent un traitement anti-reflets adapté à la longueur d'onde de la source laser.

4.4.3 Source de lumière

C'est un module commercial (Melles-Griot 06DLL307) qui se présente sous la forme d'un cylindre intégrant la diode laser, un dispositif à effet Peltier pour le contrôle en température et une optique destinée à fournir un faisceau gaussien cylindrique de diamètre 8 mm (à $1/e^2$ en puissance). Cette circularisation du faisceau, réalisée par une paire de prismes dits anamorphoseurs est nécessaire: il se trouve en effet que la cellule de Bragg choisie requiert une polarisation dans le sens de l'ouverture alors que les diodes laser fournissent un faisceau très astigmatique et elliptique polarisé selon le petit axe.

La diode laser, de type monomode et émettant à une longueur d'onde de 780 nm, est conçue pour fournir une puissance optique de 10 mW; elle est utilisée à environ 80 % de cette valeur pour augmenter sa durée de vie. Elle dispose d'une photodiode intégrée, utilisée pour l'asservissement en puissance, qui reçoit la lumière émise par la face arrière de la cavité.

L'une des caractéristiques des diodes laser est que leur longueur d'onde varie avec la température, d'une part de façon lente et continue, et d'autre part de façon brusque autour de seuils qui varient d'une diode à l'autre, ce que l'on appelle les sauts de mode. Pour combattre ces effets, le dispositif à effet Peltier intégré asservit la température à une valeur de consigne que l'on choisit éloignée des seuils de saut de mode.

La régulation en puissance et le contrôle en température de la diode sont assurés pour le moment par un instrument spécifique conçu pour un usage de laboratoire. Il est prévu de le remplacer par un module nettement plus compact mais les premières versions de ces modules ont exhibé certains défauts pouvant provoquer la destruction de la diode, composant extrêmement fragile tant au point de vue de l'électricité statique que du courant maximum qu'elle peut supporter. Cette substitution n'aura donc lieu qu'une fois que ces alimentations auront fait preuve de leur fiabilité.

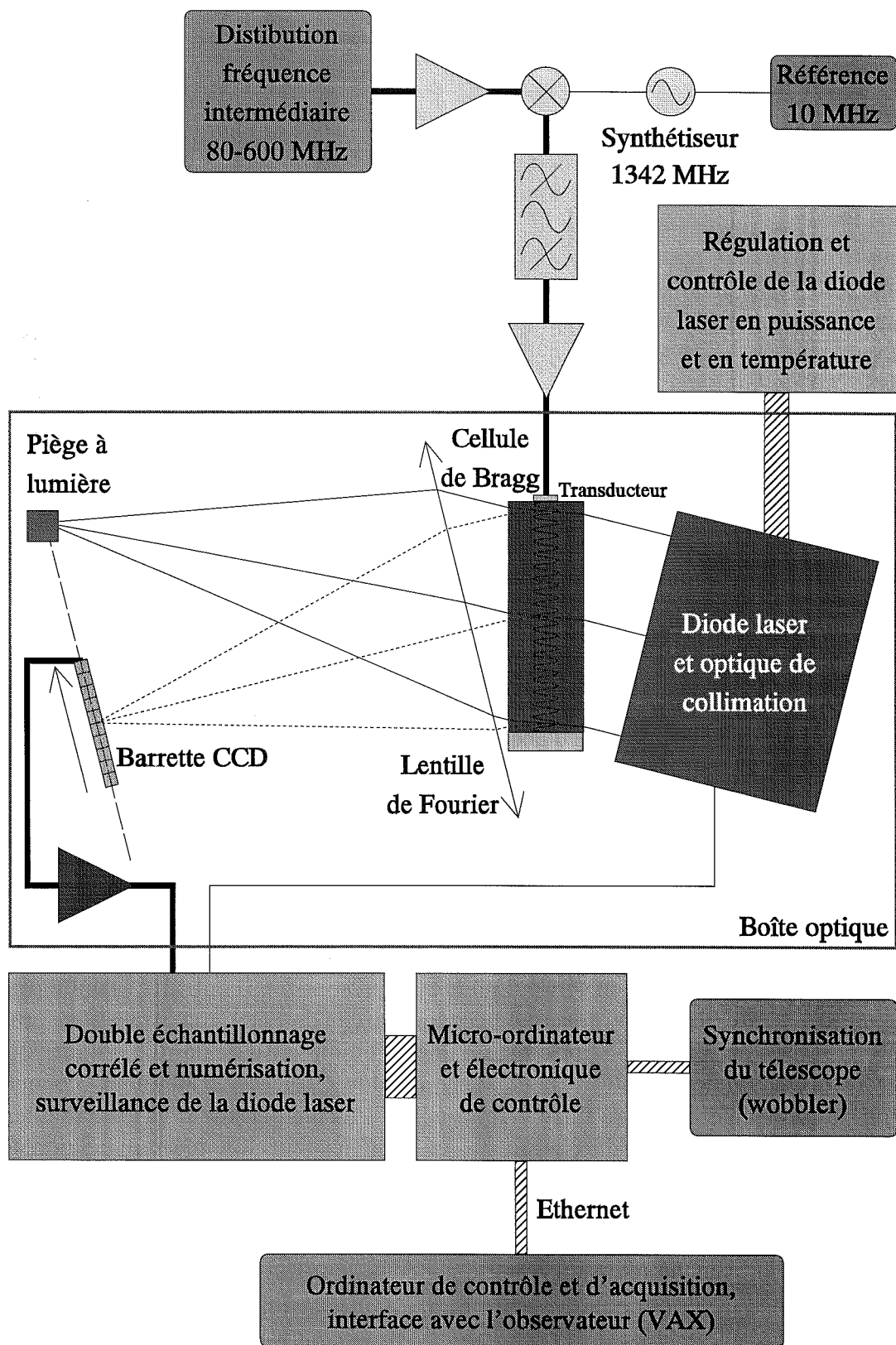


Figure 4.3: Architecture générale du spectrographe.

Enfin, le faisceau optique étant nettement plus grand que nécessaire, une fente limite l'éclairement de la cellule de Bragg à la zone d'interaction.

4.4.4 Optique et photodétecteur

Le photodétecteur est une barrette DTC (dispositif à transfert de charge, en anglais « CCD ») dont le nombre d'éléments, 1728, correspond à un standard de téléfax. Les éléments ont une surface active de $10 \times 10 \mu\text{m}$ et sont espacés de $10 \mu\text{m}$ pour une longueur totale de 17.28 mm.

Ces chiffres, conjugués avec ceux de la cellule, donnent une distance focale de 288 mm pour couvrir une bande passante de 500 MHz. La lentille est en fait un doublet pour obtenir un champ plat; étant donnée la valeur de la focale, un miroir situé à peu près à égale distance de la lentille et du photodétecteur replie le faisceau pour limiter la longueur totale du banc optique.

Dans le but de diminuer la lumière parasite, un piège à lumière est installé sur le chemin de la lumière non diffractée, à côté du miroir de renvoi.

La boîte optique est construite sur une base en acier sur laquelle sont fixés tous les composants. La base et le couvercle sont recouverts d'une peinture noire très peu réfléchissante et pratiquement inrayable, qualité appréciable quand la maintenance doit être assurée par des électroniciens qui n'ouvrent la boîte qu'exceptionnellement et qui risquent, malgré tous leurs efforts, de commettre une fausse manœuvre bien plus facilement qu'un opticien chevronné.

4.4.5 Electronique d'acquisition et de contrôle

Le fonctionnement de l'électronique d'acquisition est décrit en détail à l'appendice B, ce qui suit se limite donc à compléter quelques points et à analyser les performances.

Double échantillonnage corrélé

Le DTC est un dispositif basé sur le transistors à effet de champ à grille isolée par oxyde (MOSFET). Ces transistors se caractérisent par un bruit rose assez important, avec une fréquence de remontée comprise en général entre 10 et 100 kHz.

Mais la structure du circuit de sortie d'un DTC, qui sont tous contruits selon le principe présenté en [37, figure 9b] à ceci près que le condensateur de l'électromètre est en pratique une diode polarisée en inverse, se prête bien à une technique dite de double échantillonnage corrélé. Cette technique consiste à mémoriser, à l'aide d'un échantillonneur-bloqueur, la tension de sortie de l'électromètre entre la remise à zéro du condensateur et le flanc d'horloge qui provoquera le transfert de la charge sur ce condensateur. La tension mémorisée est ensuite soustraite de celle mesurée après le transfert. Cette technique présente plusieurs avantages:

- elle élimine le bruit de remise à zéro du condensateur qui vaut $\sqrt{kT/C}$ en tension ou \sqrt{kTC} en charge où C est la capacité totale de l'électromètre, incluant les capacités parasites. Les paramètres du DTC, dont la dispersion est importante, donnent $C \simeq 0.1 \text{ pF}$ soit 130 électrons ou $200 \mu\text{V}$ de bruit de remise à zéro. Ce résultat, auquel j'étais aussi arrivé avant de trouver cette référence, peut paraître surprenant puisque les condensateurs sont toujours considérés comme des composants sans bruit; en fait ce que réalise le transistor de remise à zéro est une modification des paramètres du spectre

de bruit du circuit, alors que l'intégrale de ce spectre reste invariable. La constante de temps du circuit varie entre quelques nanosecondes et quelques millisecondes, dans ce dernier cas on peut considérer que la charge résiduelle ne varie pas entre les deux mesures de la tension de sortie de l'électromètre, dont la différence mesure réellement la charge injectée lors du transfert du paquet d'électrons suivant.

- elle réalise un filtrage passe-haut et donc en théorie élimine le bruit rose dû au DTC. Dans le circuit réalisé ici, l'échantillonnage de la charge a lieu $1.25 \mu\text{s}$ après celui de la remise à zéro, ce qui est équivalent à un filtre de réponse sinusoïdale avec des maximums tous les multiples impairs de 400 kHz : $R(\nu) = \sin^2(\pi\nu/800000)$. Ce filtre élimine efficacement le bruit rose, y compris celui de la source de tension de remise à zéro que nous avons ignoré jusqu'ici.

Dans le cas du spectrographe acousto-optique, c'est le deuxième point qui est important car, comme nous allons le voir par la suite, la dynamique n'est pas limitée par le bruit propre du DTC. Quant au bruit de l'électromètre lui même, il peut être estimé à $10 \text{ nV.Hz}^{-1/2}$ à partir du premier terme de [37, équation 36] sachant que $g_m \simeq 10^{-3} \Omega^{-1}$. Sur une bande passante que l'on peut grossièrement estimer à 4 MHz , on obtient une valeur négligeable: $20 \mu\text{V}$ efficace, soit 10^{-5} de la pleine échelle.

La dernière source de bruit importante du DTC est le courant d'obscurité, mais la mesure avec de long temps d'intégration indique qu'il faudrait environ 10 secondes pour atteindre la saturation à la température d'utilisation. Cela signifie qu'à la fréquence de trame choisie plus loin, on récupère environ 2000 électrons par élément lu, et donc un bruit estimé de 45 électrons.

On en arriverait à la conclusion que la dynamique d'acquisition, d_{acq} charge en électrons à la saturation divisée par le bruit, est proche de 30000. En fait, d'autres sources de bruit mal maîtrisées à l'heure actuelle la limitent à 6000, valeur donnée par le fabricant lorsqu'on utilise le DTC en mode de double échantillonnage corrélé. On peut s'attendre à des progrès importants dans ce domaine puisque les derniers modèles annoncés, mais non disponibles au moment de la conception du circuit, prétendent porter cette dynamique à 10000, voire 12000 alors que la charge à la saturation varie peu.

La lumière parasite est une autre source de bruit que l'on peut considérer comme équivalente au courant d'obscurité tant qu'elle est stable. Elle est en fait nettement plus forte que ce dernier puisqu'elle saturerait le DTC en à peu près 1.5 secondes, et ce après un réglage soigneux de l'optique qui permet de la minimiser. Elle ramènerait la dynamique d'acquisition à une valeur plus proche de 10000, qui n'est pas complètement négligeable par rapport au bruit électronique, la somme des deux donnant $d_{\text{acq}} \simeq 5000$.

Limitations dûes au DTC

La charge à la saturation Q_{sat} de tous les DTC disponibles à l'heure actuelle est de l'ordre du million de photoélectrons par élément, environ $1.25 \cdot 10^6$ pour le modèle choisi. Le processus de création de ces photoélectrons obéit à une statistique Poissonnienne qui possède un bruit incompressible égal à la racine de leur nombre.

Lorsqu'on considère un spectrographe dont la bande équivalente de fluctuation est de l'ordre de 1.7 MHz , comme dans notre cas, on voit que pour que le bruit de photons n'implique pas une augmentation de plus de 10 % de la variance, il faut détecter plus de $17 \cdot 10^6$ photoélectrons par seconde.

D'un autre côté, la charge maximale par seconde que l'on peut créer dans chaque élément est limitée par la puissance que l'on peut fournir à la cellule de Bragg sans que l'échauffement produise des distortions optiques mesurables, en pratique quelques dizaines de milliwatts. Le fait que l'on soit limité par ce problème n'est pas étonnant si l'on considère les énormes densités de puissance que l'on peut atteindre [36, page 301]. L'amplificateur qui pilote la cellule est spécifié en fonction de ces paramètres et non de limites sur leur faisabilité, un modèle plus puissant risquant de détruire la cellule.

Au vu des caractéristiques obtenues dans la partie optique, la fréquence de trame de la CCD a été fixée à $f = 64$ Hz; on pourrait la pousser à 80 Hz sans rien changer mais le premier effet serait d'augmenter le bruit de lecture, et ne présente que l'avantage d'accepter un niveau de puissance plus élevé dans une partie de la bande (une « bosse ») sans saturation. Dans ce cas la dynamique d'un élément de la CCD vaut $64 \times 1.25 \cdot 10^6 / 17 \cdot 10^6 \simeq 4.7$, c'est insuffisant mais on peut profiter du fait que la tache focale couvre presque 4 éléments pour les sommer et augmenter ce chiffre. Les éléments sont pour le moment sommés par paires, l'observateur voit donc le spectrographe comme un instrument délivrant 864 canaux espacés de 585 kHz. Les éléments voisins sont tellement corrélés que cette procédure multiplie la dynamique par un facteur proche de 2, alors que la bande de fluctuation n'augmente que de 100 kHz. Les canaux sommés sont encore fortement corrélés, en fait l'échantillonnage résultant est quasiment parfait du point de vue du critère de Nyquist.

Le dynamique d'acquisition du DTC dû à son bruit propre devient dans ce cas $\sqrt{2}d_{\text{acq}} \simeq 7000$. Pour ne pas rajouter de bruit d'après nos critères, la limitation correspondante à la dynamique est $\sqrt{2f/10B_{\text{ac}}} d_{\text{acq}} \simeq 13.3$. Cette valeur est supérieure de 50 % à la limite sur le bruit statistique, mais alors que l'accroissement du temps de pose requis causé par ce dernier est inversement proportionnel au nombre de photons recueillis, celui dû au bruit fixe varie comme le carré de cet inverse. Il y aura donc une gamme de puissance dans laquelle le bruit croîtra progressivement avant d'augmenter rapidement. Sommer les éléments du DTC par groupes de plus de 2 n'aurait amélioré que très peu la dynamique, on peut considérer le choix fait ici comme le meilleur compromis.

Si on avait pris un critère moins sévère sur l'excès de bruit tolérable, on aurait pu être limité par le bruit propre du DTC; c'est le cas général pour les spectrographes à haute résolution. Le point où les deux excès de bruit sont égaux dans notre cas se situe vers 5 % de la pleine échelle et correspond à un facteur de perte en temps d'intégration de 1.45, inacceptable en pratique.

4.4.6 Linéarité

La linéarité de l'interaction acousto-optique ne fait pas de doute, on est en pratique limité par l'échauffement de la cellule avant tout autre phénomène. Les DTC sont eux aussi des composants extrêmement linéaires. Nous reviendrons plus tard sur les problèmes de linéarité causés par l'amplificateur de puissance lors de la comparaison avec les batteries de filtres. Il reste un composant critique: le convertisseur analogique-numérique. Les graphiques montrés par les fabricants montrent souvent la qualité de la réponse par la transformée de Fourier d'un signal monochromatique d'amplitude proche du maximum possible. Ces graphiques sont très sensibles à l'erreur de linéarité globale, qui ferait apparaître des harmoniques, mais sont peu sensibles à l'erreur de linéarité différentielle. Or, comme le signal pour un élément donné du DTC varie très peu, son incertitude statistique relative étant environ 0.6 %, les problèmes de linéarité différentielle du convertisseur peuvent être la plus importante source d'erreur du spectrographe.

Lorsque le mot binaire de sortie d'un convertisseur analogique-numérique idéal a la valeur n , cela signifie que la tension d'entrée V est comprise entre deux seuils notés V_n et V_{n+1} . Pour un convertisseur possédant N points de résolution, on considérera que la valeur de sortie est comprise entre 0 et $N-1$ inclus et que les seuils V_n ont tous une valeur finie à l'exception de $V_0 = -\infty$ et $V_N = +\infty$. Les autres seuils sont théoriquement distribués régulièrement entre V_1 et V_{N-1} avec un intervalle entre deux seuils $V_u = \frac{V_{N-1}-V_1}{N-2}$ et le convertisseur unipolaire parfait sans tension d'erreur vérifie $V_n = (n - \frac{1}{2})V_u$ pour $1 \leq n \leq N-1$.

Si un tel convertisseur échantillonne un signal aléatoire de densité de probabilité $P(V)$ au point V , la tension moyenne mesurée sera:

$$V_m = \sum_{n=0}^{N-1} nV_u \int_{V_n}^{V_{n+1}} P(V) dV$$

où $P(V)$ est la densité de probabilité de la tension de sortie du DTC autour de la valeur V . On peut naturellement faire l'hypothèse que la distribution $P(V)$ tend vers zéro plus rapidement que toute puissance de $1/V$ lorsque V tend vers $\pm\infty$.

On supposera aussi que lorsque la tension moyenne $\bar{V} = \int_{-\infty}^{+\infty} VP(V) dV$ varie, la densité de probabilité garde la même forme, on peut donc se ramener à une densité normalisée à la valeur moyenne et réécrire l'expression précédente sous la forme:

$$V_m = \sum_{n=0}^{N-1} nV_u \int_{u_n}^{u_{n+1}} p(u) du$$

avec $u_n = \frac{V_n}{\bar{V}}$ et $p(u) = \bar{V}P(V)$. Lorsque la tension moyenne varie, ce sont maintenant les bornes de l'intégrale qui changent et non plus la densité de probabilité à intégrer. Calculons la dérivée de la tension mesurée par rapport à \bar{V} :

$$\frac{dV_m}{d\bar{V}} = \sum_{n=0}^{N-1} nV_u \left(p(u_{n+1}) \frac{du_{n+1}}{d\bar{V}} - p(u_n) \frac{du_n}{d\bar{V}} \right)$$

qui se simplifie considérablement:

$$\frac{dV_m}{d\bar{V}} = \frac{V_u}{\bar{V}} \sum_{n=1}^{N-1} V_n p(V_n)$$

si l'on tient compte des hypothèses faites sur $P(V)$.

Idéalement, cette dérivée devrait être une constante. Deux causes font qu'elles ne l'est pas exactement; la première est le fait que la somme n'est pas rigoureusement constante même sous l'hypothèse d'un convertisseur parfait, la deuxième est que les seuils V_n sont loin d'être régulièrement répartis.

Ce dernier point est dû au fait que les convertisseurs rapides utilisant la méthode des approximations successives que l'on est quasiment obligé d'utiliser dans notre cas sont contruits autour de convertisseurs digitaux-analogiques basés sur des échelles résistives de type R-2R. Dans ce type de circuit, l'appariement des résistances est critique et les problèmes liés aux dérives thermiques et au vieillissement des composants, bien qu'atténués par l'intégration des éléments sur la même puce, font que l'erreur d'espacement des seuils du convertisseur est souvent donnée comme pouvant atteindre 0.75 voire 0.9 V_u . Une autre source de problèmes vient du comparateur qui décide à chaque itération si la tension d'entrée est supérieure à un seuil qui dépend du résultat des comparaisons précédentes. On demande à ce comparateur

d'exceptionnelles qualités dynamiques: en particulier un produit gain bande assez faramineux (de l'ordre de 10 GHz) et de récupérer très vite d'un état de saturation. Enfin, l'effet du temps de stabilisation du convertisseur numérique-analogique interne et l'impédance de sortie de l'échantillonneur bloqueur à haute fréquence sont d'autres paramètres pouvant affecter sérieusement la linéarité différentielle.

Une solution à ce problème consiste à utiliser des convertisseurs de type « flash », qui comportent autant de comparateurs que de seuils et présentent l'avantage d'une plus grande rapidité. Ce genre de convertisseur est en pratique limité à 10 bits de précision, ce qui demande déjà plus de 1000 comparateurs de précision sur une puce. Lorsqu'une meilleure résolution est requise, on a recours à une technique mixte, dans laquelle une première échelle de comparateurs détermine les bits les plus significatifs, qui sont aussi dirigés vers l'entrée d'un convertisseur numérique analogique (CNA). On peut alors déterminer un résidu correspondant à la différence entre la tension à mesurer et la tension fournie par le CNA, qui est convertie par une autre échelle de comparateurs. Ceci peut éventuellement être répété, et on pourrait plutôt appeler cette méthode conversion par approximation successive avec détermination simultanée de plusieurs bits. Cette technique présente cependant aussi l'inconvénient d'erreurs de linéarité différentielle importantes autour des seuils de la première échelle de comparateurs.

La solution proposée ici à ce problème est une variation sur une technique souvent utilisée avec les convertisseurs analogique-numérique audio appelée « dither » en anglais et qui sera ici traduit, faute de mieux, par « brouillage ».

Par exemple, l'analyseur de spectre utilisé pour la mise au point des détecteurs continu, qui échantillonne à 2 fois la fréquence nécessaire le signal d'entrée filtré auquel on a rajouté un bruit généré par un oscillateur « sale » compris entre f et $2f$. Le filtrage numérique ensuite réalisé élimine le signal de brouillage.

On ne peut pas utiliser directement ce genre d'astuce dans le cas de la lecture d'un DTC car cela revient à rajouter un bruit individuel supérieur au bruit du signal sur chaque élément. Par contre le signal de sortie du DTC, que l'on peut considérer comme un intégrateur quasi parfait, est proportionnel au temps d'intégration. C'est en faisant varier ce dernier que l'on peut moduler de façon globale le signal à numériser. Les calculs théoriques pour une erreur de linéarité différentielle de $0.5V_u$ au point de poids 512, encore dans la gamme utilisables dans notre cas sont présentés à la figure 4.4. L'erreur maximale qui est presque de 5 % dans le cas d'une lecture de durée fixe, est divisée par 10 et devient négligeable dans le budget des erreurs d'étalonnage.

Il faut souligner que cette technique, comme elle résulte en une multiplication de la tension moyenne sortie du DTC par une constante comprise entre 0.9 et 1.1, n'est pas utilisable si le signal tombe en dessous de quelques dizaines de comptes du convertisseur, valeur pour laquelle la tension à convertir ne se répartit plus que sur quelques seuils. C'est sans importance car on a vu qu'on est déjà limité par la statistique de bruit de photons du DTC.

En pratique, cette modulation est réalisée en reprogrammant, chaque fois que l'interruption de fin de lecture du DTC est reconnue, le registre commandant la durée d'intégration. Le mot écrit dans ce registre est la somme d'une valeur moyenne et du résultat d'un générateur aléatoire.

Malgré de nombreuses expériences, il a été impossible de trouver un exemple démontrant une amélioration ou une dégradation des caractéristiques du spectrographe due à cette astuce. Elle a donc été mise en place pour des raisons théoriques et parce que le fait que la technique de brouillage soit utilisée dans de nombreux instruments professionnels implique

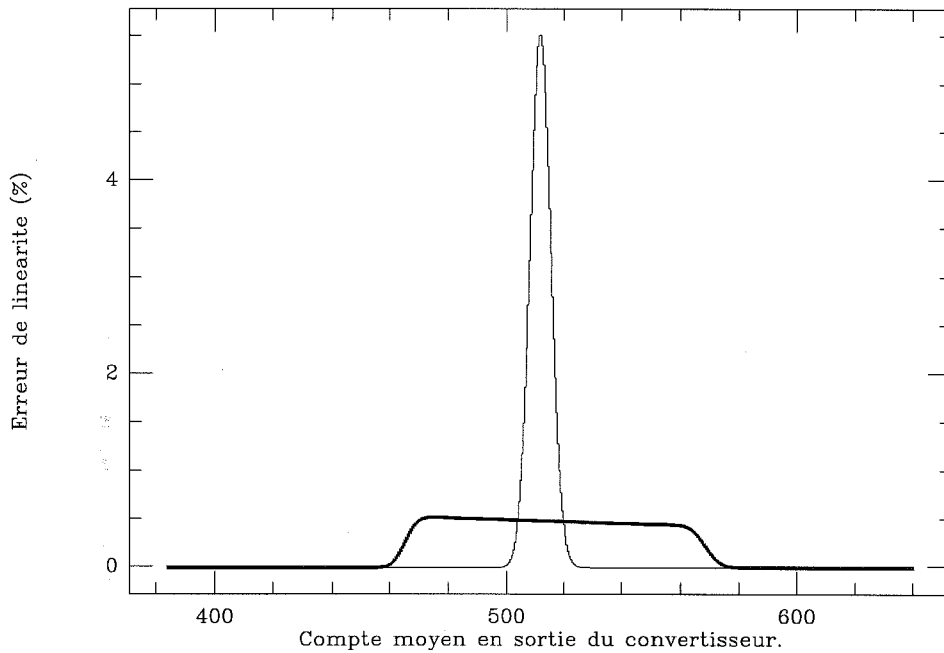


Figure 4.4: Erreur de linéarité sans (trait fin) et avec brouillage (trait gras)

qu'elle a une utilité.

4.4.7 Processeur de fréquence intermédiaire

Son rôle est de transposer la bande de fréquence intermédiaire fournie par le système de distribution du radiotélescope dans vers la bande de fréquence de la cellule de Bragg et d'amplifier le signal jusqu'au niveau de puissance requis. L'apparition de circuits intégrés monolithiques a considérablement diminué le coût et la complexité de ce genre de fonction, qui était l'apanage de spécialistes il y a seulement quelques années.

Pour des raisons historiques, la bande de fréquence intermédiaire du télescope est centrée à 342 MHz. Une boucle à verrouillage de phase, dont la référence est l'horloge maîtresse du télescope, asservit un oscillateur contrôlé en tension à la fréquence de 1342 MHz. Ce signal est mélangé avec la fréquence intermédiaire du télescope et ensuite filtré pour sélectionner une bande de 500 MHz centrée en 1 GHz. C'est là que la seule petite difficulté apparaît; en effet, comme la bande de la cellule de Bragg à mi-puissance est à peine supérieure à 500 MHz et pour ne pas compromettre la dynamique, la largeur du filtre à mi-puissance vaut 550 MHz. Le résultat est que les fuites de l'oscillateur local, situé seulement 67 MHz (5 % en valeur relative) au dessus de la fréquence de coupure du filtre, sont excessives. Il a fallu rajouter entre deux amplificateurs monolithiques un réjecteur, simplement un condensateur et une inductance en série entre la piste de signal et la masse. Pour éliminer ce bricolage, qui atténue un peu le signal en bord de bande, un filtre interdigital extrêmement sélectif a été commandé mais les délais de livraison et d'autres contraintes n'ont pas encore permis de l'installer.

Un amplificateur de puissance commercial amène le signal transposé et filtré au niveau requis par la cellule de Bragg, et même plus puisqu'un atténuateur de 6 dB est placé entre

les deux pour limiter l'effet des ondes stationnaires.

4.4.8 Processeur local

C'est une carte standard (GESPACE SBS-6) avec un processeur Motorola 68000 tournant sous le système d'exploitation temps-réel OS-9. Les signaux de synchronisation du télescope, utilisés avec le secondaire oscillant, sont connectés à des lignes d'interruptions de cette carte.

Sa tâche essentielle, lire le résultat des conversions mémorisées dans les tampons de type premier-entré premier-sorti (« FIFO ») de la carte d'interface, et les accumuler dans la mémoire locale pendant la durée d'une pose en fonction du mode d'observation, est tellement simple qu'elle ne mérite pas de plus amples explications. Signalons seulement qu'il a fallu optimiser sérieusement, en l'écrivant en langage d'assemblage, la boucle la plus critique, celle de lecture des tampons. Après optimisation, cette boucle occupe environ 30 % du temps du processeur.

4.4.9 Interface avec l'ordinateur de contrôle du télescope

La figure 4.3 est en réalité un peu en avance sur son temps. En effet, on utilise pour le moment le port parallèle de la carte microprocesseur pour communiquer avec un module situé dans une baie CAMAC. C'est une solution simple, qui a été utilisée depuis l'installation du spectrographe et donne satisfaction depuis que les problèmes de l'installation de plusieurs microprocesseurs sur le même fond de panier CAMAC ont été résolus.

Cependant, à la lumière de l'expérience acquise avec les nouveaux corrélateurs livrés en Février 1992 et installés au télescope en Mai, il paraît préférable de s'orienter vers une solution réseau local type Ethernet. Après quelques problèmes de démarrage bien compréhensibles, la fiabilité des nouveaux corrélateurs sur le plan logiciel paraît, 6 mois après leur mise en service, excellente. Comme le standard CAMAC est complètement obsolète, les composants critiques de nombreux modules utilisés au télescope n'étant plus fabriqués, et que les éléments nécessaires pour connecter le spectrographe acousto-optique à un réseau local sont disponibles dans le commerce à un prix raisonnable, il semble justifié de réaliser cet investissement. Cette amélioration ne devrait en pratique entraîner strictement aucun changement pour l'observateur.

4.5 Caractéristiques mesurées

Les paragraphes qui suivent présentent les caractéristiques mesurées au télescope, sauf exception après l'installation définitive en Mars 1990.

4.5.1 Comparaison avec les batteries de filtres

Lors de la première installation du spectrographe en Octobre 1989, une intégration longue a été effectuée sur la source IRC+10216 pour rechercher le radical CP. Lors de cette observation, les deux batteries de filtres et le spectrographe acousto-optique analysaient le même signal. Bien qu'il s'agisse là d'un cas de figure où l'on utiliserait de préférence le secondaire vibrant, l'observation fut réalisée en mode « position-switch » avec des cycles d'une minute dans le but de rechercher les éventuels défauts des spectrographes.

26; 2 IRC+10216 238 IRAM-30M-B60 O: 8-OCT-1989 R: 8-OCT-1989
RA: 9:45:14.800 DEC: 13:30:40.00 (1950.0) Offs: 0.0 0.0 Eq
Unknown Tau: 0.4783 Tsys: 1575. Time: 140.0 El: 62.39
N: 256 IO: 128.5 VO: -27.00 Dv: 2.516 LSR
FO: 238304.519 Df: -2.000 Fi: 246167.384
B ef: 0.9000 F ef: 0.9000 G im: 0.2500

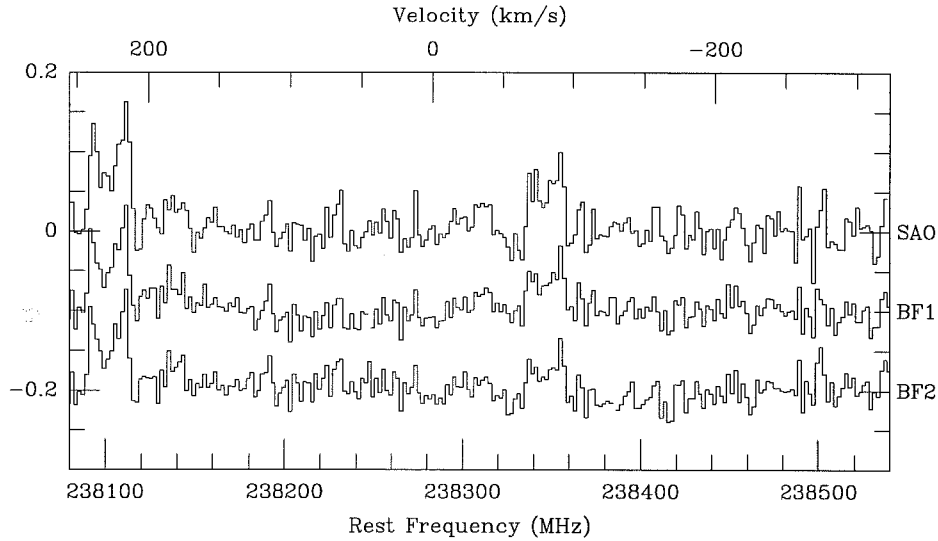


Figure 4.5: Comparaison du spectrographe acousto-optique (SAO) avec les batteries de filtres (BF1 et BF2)

Les trois spectres, lissés à 2 MHz de résolution pour limiter l'effet des différences d'échantillonnage entre les différents spectromètres, sont montrés en figure 4.5 après un temps d'intégration de 140 minutes. Ce spectre a été publié [15]; sur le plan technique, l'accord entre les 3 instruments démontre le bon fonctionnement des trois spectrographes.

On remarquera une différence d'étalonnage sur ces spectres, ceci est dû à une saturation de l'amplificateur de puissance. Pour l'installation définitive en Mars 1990, cet amplificateur a été révisé par le fabricant pour augmenter son niveau de compression de 2 dB, soit 60 %, ce qui a rendu ces problèmes négligeables, la différence mesurée maintenant entre le spectrographe acousto-optique et les batteries de filtres n'excédant pas 3 %; diminuer un peu l'atténuation entre l'amplificateur et la cellule permettrait de l'améliorer encore, mais il faudrait refaire toute une série d'essais pour vérifier que la stabilité n'est pas dégradée par l'augmentation des ondes stationnaires. Signalons aussi que pour ce premier essai au télescope, la lentille de Fourier n'avait pas la même focale que la version définitive, et que la bande totale était donc un peu plus faible, 468 MHz au lieu de 505.

4.5.2 Filtre équivalent

La forme du filtre est présentée à la figure 4.6, elle a été obtenue en acquérant avec un canal du spectrographe le signal monochromatique fourni par un synthétiseur. Cette courbe est le résultat de 101 mesures obtenues en variant la fréquence du signal par pas de 200 kHz sur une bande de 20 MHz. On peut en déduire les paramètres de largeur à -3, -10 et -20 dB qui sont respectivement 1.1, 2.1 et 5.6 MHz. On voit clairement qu'en dessous de -10 dB, et surtout en dessous de -15 dB, la forme du spectre n'est plus du une gaussienne, qui se traduirait par une parabole sur l'échelle logarithmique employée. Si une source laser

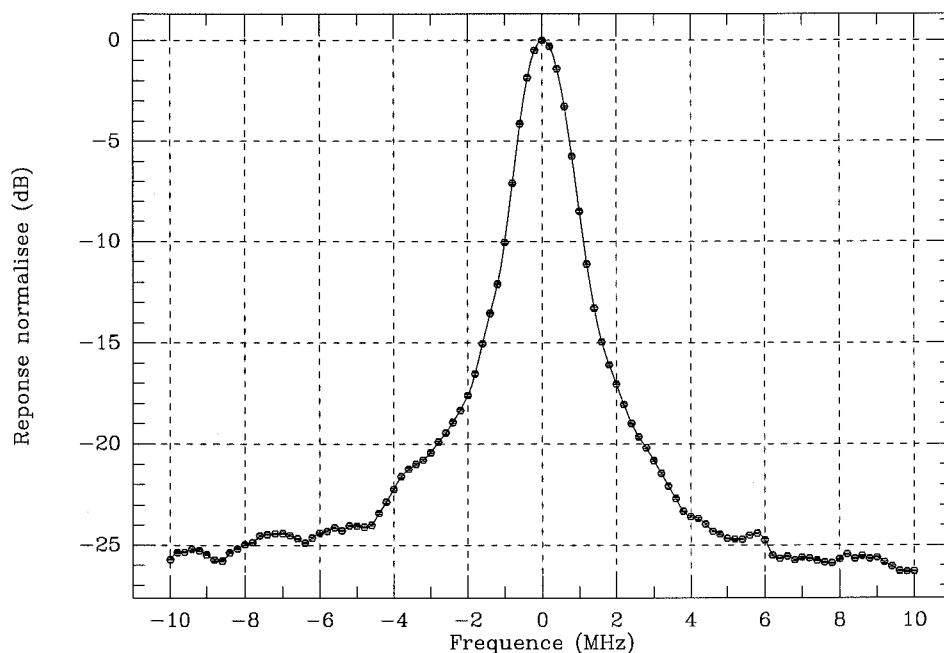


Figure 4.6: Forme du filtre équivalent à un canal du spectrographe

donnant une largeur mieux adaptée à la cellule avait été disponible, elle aurait peut-être permis de diminuer cet réponse parasite au prix d'une très faible augmentation de la largeur à mi-hauteur.

Pour étudier l'effet de cette réponse en pratique, il faut prendre une raie astronomique forte, par exemple, la deuxième transition de CO à 230.538 GHz dans une enveloppe d'étoile carbonée comme IRC+10216; cette raie est montrée en figure 4.7. Elle présente une forme caractéristique de raie saturée dans cette source, avec des flancs très raides, idéaux pour étudier les défauts de la réponse en fréquence.

Un agrandissement de ce spectre permettant une comparaison valide est donné en figure 4.8. La différence est claire mais n'est que très exceptionnellement gênante, l'intégrale de chaque aile vaut à peu près 1 % de l'intégrale de la raie; c'est le cas le plus défavorable puisque la largeur de chaque aile est à peu près égale à la largeur de la raie. Il faut cependant avertir les astronomes qui étudient les ailes de raie qu'il doivent faire attention aux possibles réponses parasites du spectrographe.

L'origine de ce défaut n'est pas bien comprise, aucun phénomène optique ne permettant d'expliquer de façon satisfaisante la forme observée, qui paraît à peu près triangulaire. Il s'agit peut être d'un phénomène de diffusion dans le cristal qui, rappelons-le, n'est pas parfaitement transparent.

On remarquera au passage la très grande similitude des détails du spectre entre les trois spectrographes, prouvant clairement que leur contribution au bruit est très faible.

4.5.3 Variance d'Allan

La figure 4.9 présente la variance d'Allan d'un canal du spectrographe et d'un canal des batteries de filtres de 1 MHz acquis simultanément et à la même fréquence dans la bande.

4; 1 IRC+10216 12CO(2-1) IRAM-30M-B60 O: 21-JUL-1990 R: 21-JUL-1990
RA: 9:45:14.800 DEC: 13:30:40.00 (1950.0) Offs: 0.0 0.0 Eq
Unknown Tau: 0.6741 Tsys: 1496. Time: 40.00 El: 57.03
N: 864 IO: 436.9 VO: -27.00 Dv: -0.7597 LSR
FO: 230537.990 Df: 0.5842 Fi: 238401.282
B ef: 0.9000 F ef: 0.9000 G im: 0.2500
1185, 1187, 1189,

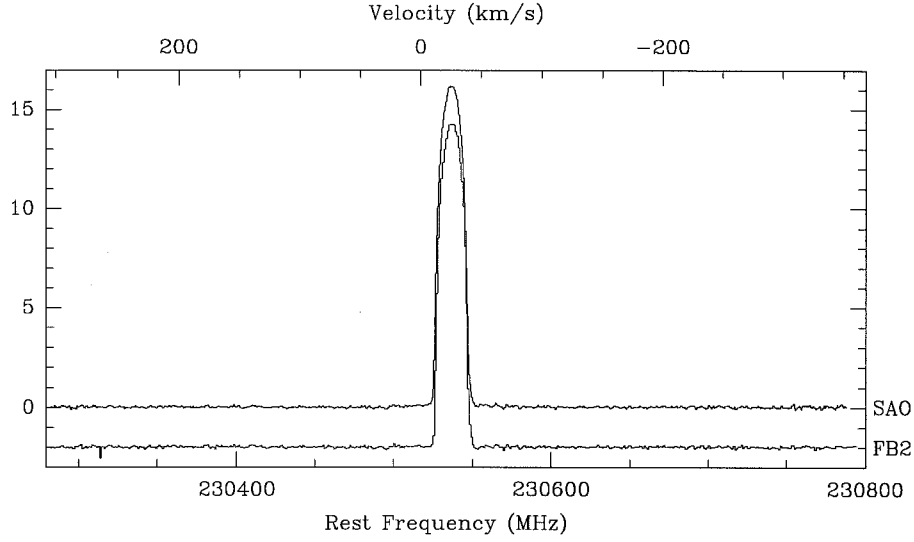


Figure 4.7: Raie du monoxyde de carbone à 230.538 GHz dans IRC+10216

4; 1 IRC+10216 12CO(2-1) IRAM-30M-B60 O: 21-JUL-1990 R: 21-JUL-1990
RA: 9:45:14.800 DEC: 13:30:40.00 (1950.0) Offs: 0.0 0.0 Eq
Unknown Tau: 0.6741 Tsys: 1496. Time: 40.00 El: 57.03
N: 864 IO: 436.9 VO: -27.00 Dv: -0.7597 LSR
FO: 230537.990 Df: 0.5842 Fi: 238401.282
B ef: 0.9000 F ef: 0.9000 G im: 0.2500
1185, 1187, 1189,

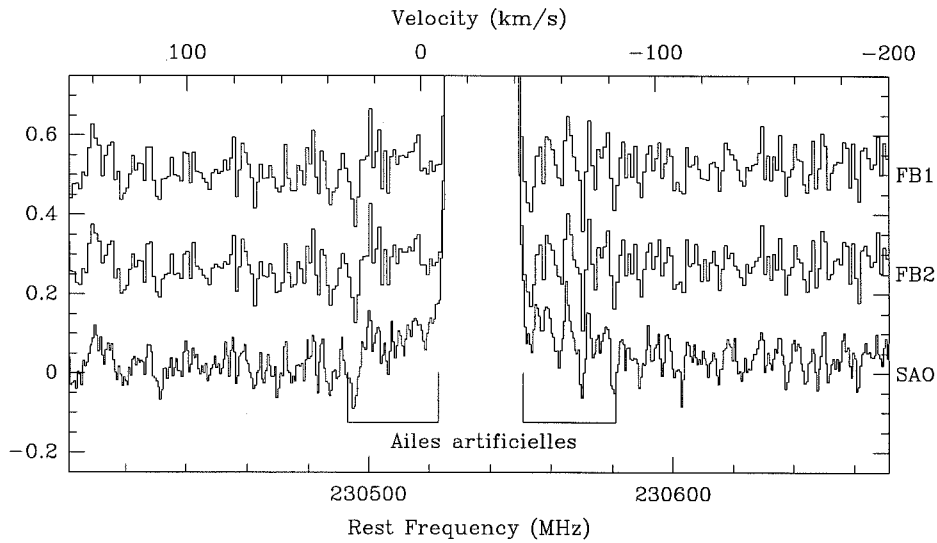


Figure 4.8: Agrandissement de la figure précédente montrant les pieds de la réponse en fréquence.

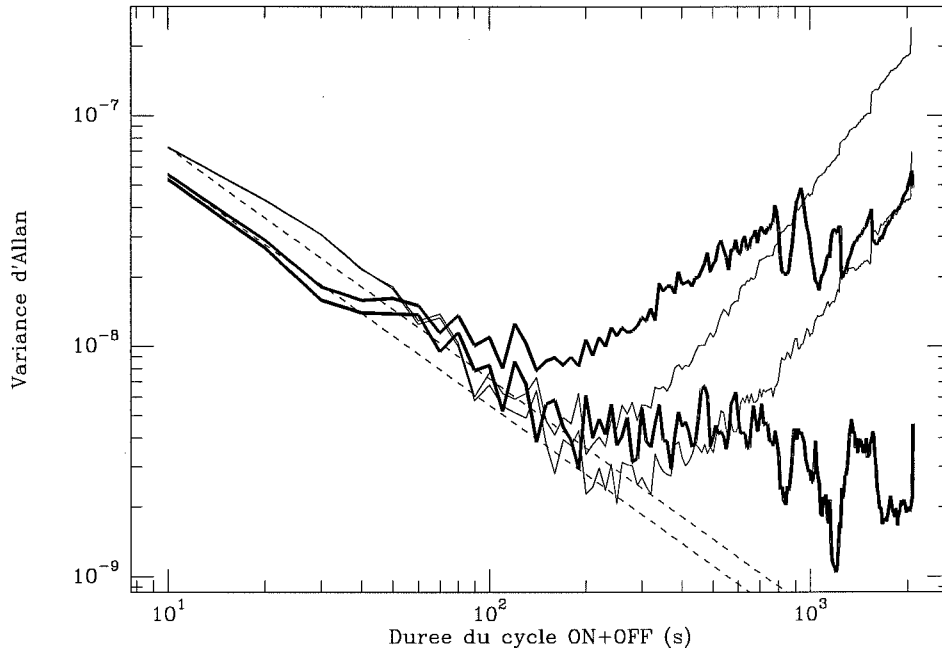


Figure 4.9: Variance d'Allan du spectrographe (trait gras) et des filtres (trait fin) avec et sans correction de gain

Les courbes inférieures ont été corrigées du gain, c'est à dire qu'avant de calculer la variance, les échantillons ont été divisés par une estimation de la puissance totale obtenue comme la somme de tous les canaux divisés par leur gain respectif. Dans le cas du spectrographe acousto-optique, la différence est importante et montre que, bien que la stabilité absolue du spectrographe soit assez limitée, le résultat ne sera en pratique qu'une ligne de base de degré zéro pour des cycles de durée le l'ordre de la minute, utilisés en pratique en mode « position-switch ». La dégradation devient par contre très nette au delà de 3 minutes, ce qui est parfaitement acceptable. La valeur de la band de fluctuation déduite de ces courbes est 1.8 MHz comme l'indique le trait pointillé. Dans le cas des filtres, l'application de la correction de gain ne fait apparaître des différences qu'au delà de 100 secondes et permettrait de pousser la durée des cycles jusqu'à 5 minutes. Le fait que la variance des filtres se situe au dessus de la courbe pointillée, correspondant à une bande de fluctuation de 1.4 MHz pour redescendre ensuite et même passer un peu en dessous est pour le moment incompréhensible.

Vers les longues durées, il est intéressant de constater qu'après correction du gain le spectrographe n'exhibe, contrairement aux filtres, absolument aucune dérive, qui se traduirait par une pente de degré 1 sur la figure 4.9. Le fait que cette variance soit presque horizontale indique qu'un problème de bruit rose est la limitation à la stabilité. Rappelons ici que la répétition d'un cycle d'observation, qui correspondrait à calculer la variance d'Allan sur des cycles de poids $\{-1, 1, \dots, -1, 1\}$ possède un effet de filtrage qui élimine plus efficacement le bruit rose que les dérives.

En résumé, la dégradation du spectrographe démarre pour des cycles un peu pus courts que pour les filtres mais elle est très progressive. Dans les deux cas, les caractéristiques obtenues pour les durées de cycles utilisées en pratique sont très proches.

Signalons que nous nous sommes rendus compte récemment du fait que la dérive glo-

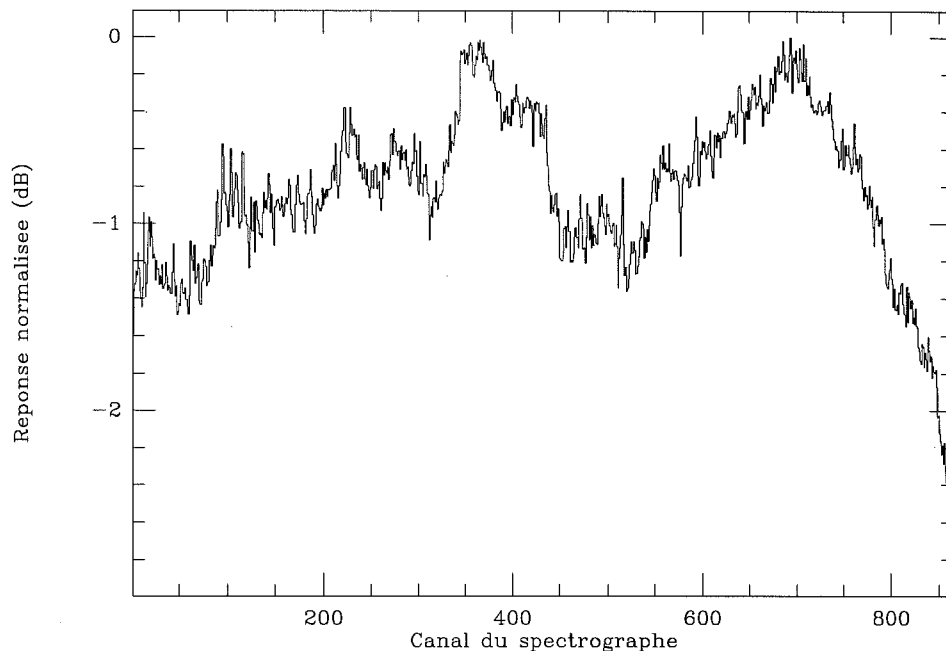


Figure 4.10: Bande passante sur le premier récepteur à 1.3 mm

bale du spectrographe provenait surtout de la piètre qualité du potentiomètre de réglage de puissance de la diode laser. Après l'avoir remplacé par un composant de meilleure qualité, la stabilité de la source de lumière, mesurée par le convertisseur tension-fréquence installé dans l'électronique d'acquisition, s'est trouvée améliorée d'un facteur 100 sur quelques heures. Nous n'avons malheureusement depuis lors pas trouvé le temps d'effectuer nouvelles mesures de la variance d'Allan; la stabilité absolue devrait s'en trouver améliorée mais l'effet devrait être minime après correction du gain.

4.5.4 Ondulation de la bande

La dynamique d'acquisition du spectrographe est, on l'a vu, assez faible à cause du bruit statistique. La dynamique utilisable lors des observations est diminuée d'autant plus que l'ondulation dans la bande sera forte.

Dans le but de diminuer cette ondulation, et après avoir analysé la bande passante des récepteurs disponibles. Il est apparu impossible de la compenser dans tous les cas. Par contre, les deux récepteurs 1.3 mm et le récepteur 2 mm se sont révélés posséder des bandes passantes assez similaires. Un circuit d'égalisation a donc été conçu et installé dans le processeur de fréquence intermédiaire, dans la partie basse fréquence avant le mélangeur puisque la bande relative plus large exigeait une sélectivité moindre.

La bande obtenue sur un récepteur est montrée à la figure 4.10 dans un cas favorable. Elle dépend en effet assez fortement de la fréquence et du réglage, simple ou double bande, du récepteur. Le cas le plus fréquent est cependant une diminution plus forte de la puissance vers le côté gauche de la figure, qui ne modifie pas l'ondulation totale. Il est difficile d'améliorer la réponse des derniers canaux sur le côté droit, qui correspondent à la fréquence acoustique la plus élevée et où la perte d'efficacité de la cellule s'ajoute à une diminution du gain du

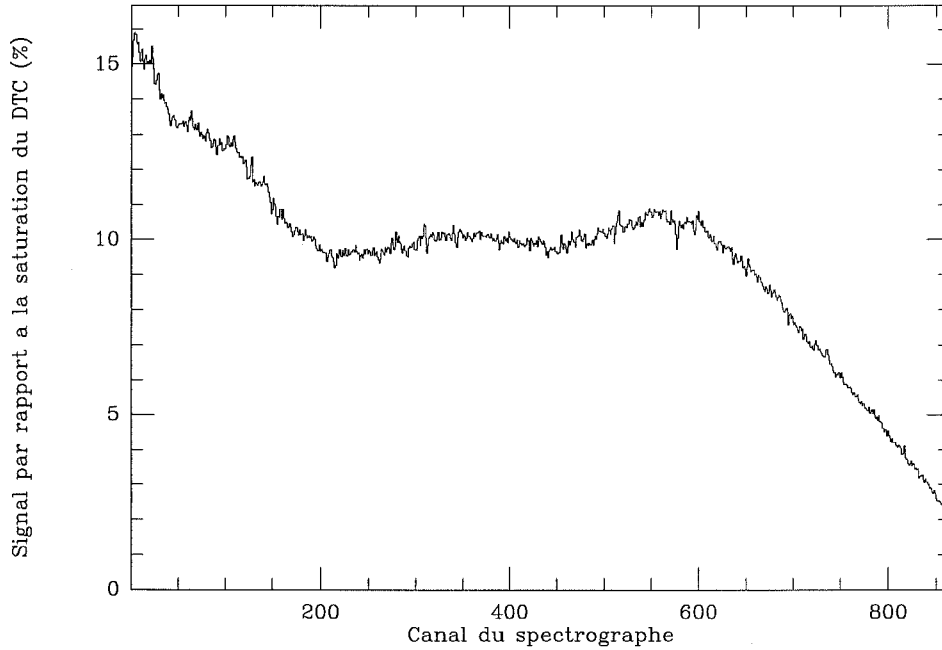


Figure 4.11: Forme de bande utilisée pour mesurer la dynamique

dernier amplificateur, à un début de coupure du filtre pass-bande et aux premiers effets de celui qui rejette l'oscillateur local.

Les variations de gain entre canaux voisins paraissent importantes, elles sont dûes au DTC et restent dans la gamme des ± 5 à $\pm 15\%$ spécifiée spécifiés par le fabricant. Elles sont de toutes façons corrigées par l'étalonnage.

4.5.5 Dynamique

Pour évaluer la dynamique, un signal à assez faible niveau et avec une forte ondulation dans la bande montré à la figure 4.11 a été envoyé au spectrographe. Cette figure donne au passage une idée de l'ondulation initiale que compense l'égaliseur, la source utilisée générant un bruit pratiquement achromatique à l'entrée du spectrographe.

Ensuite, après avoir acquis 200 spectres de 5 secondes chacun, on a calculé la variance d'Allan de chaque canal pour un cycle de 10 secondes, le plus court possible. On peut alors tracer les points donnant la bande de fluctuation équivalente en fonction du niveau, c'est ce qui est montré à la figure 4.12. Il est clair, au vu de la dispersion, que la mesure précise de la largeur d'autocorrélation requiert un très grand nombre d'échantillons et qu'il faut prendre les résultats avec beaucoup de précautions. Néanmoins, cette courbe peut être décrite par trois paramètres sous la forme:

$$y(x) = a_0(1 + a_1/x + a_2/x^2)$$

pour x variant entre 0 et 1 à la saturation. Dans ce cas, le paramètre a_1 vaut entre 0.01 et 0.012, ce qui correspond à une dynamique de 9. Quant à a_2 , il est un peu plus fort que prévu, mais il est très sensible à quelques points isolés, éliminer le point le plus élevé sur la figure 4.12 diminue sa valeur de 7%. On en déduit une dynamique d'acquisition plus

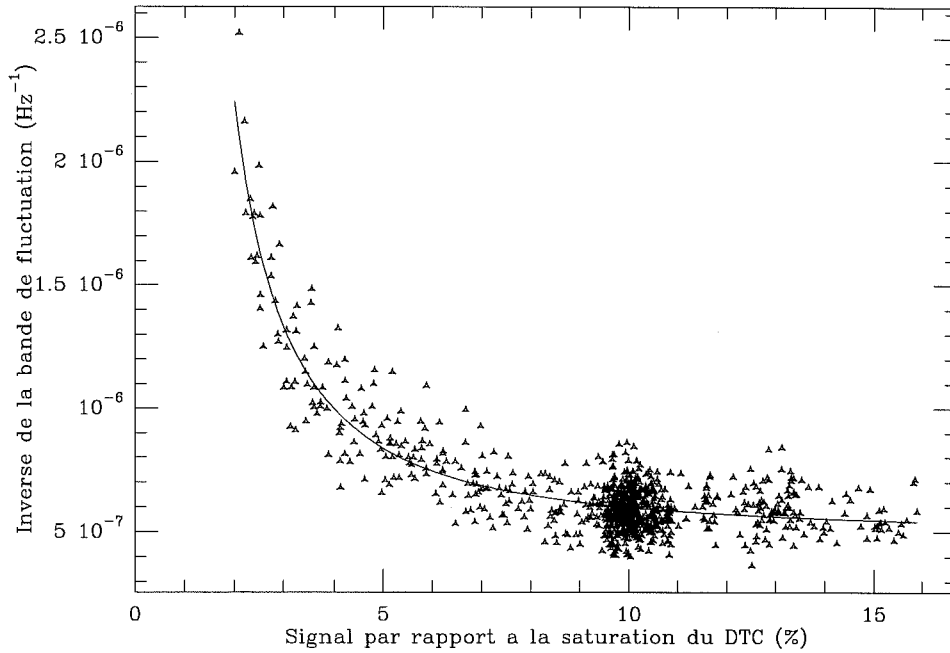


Figure 4.12: Inverse de la bande de fluctuation en fonction du niveau

proche de 4000. Il ne faut pas s'en étonner car la dispersion sur les caractéristiques des DTC est forte, le facteur d_{acq} minimum garanti par le fabricant valant seulement 3000, la moitié de la valeur annoncée comme « typique ». Nous en tirerons comme conclusion que la dynamique d'acquisition radioastronomique est comprise entre 8 et 10. Une fois divisée par l'ondulation dans la bande, et si l'on tient compte du fait que les pointes de la bande passante ne dépassent pas 85 % à cause de la saturation de l'amplificateur de puissance et ne peuvent en aucun cas dépasser 90 % à cause du brouillage, la dynamique réellement obtenue est seulement 4, cela s'est révélé utilisable en pratique mais c'est tout juste à la limite de l'acceptable, en particulier cela ne laisse pratiquement aucune marge de sécurité.

4.5.6 Etalonnage en fréquence

L'échelle en fréquence du SAO est susceptible de varier pour un certain nombre de raisons. Il a par exemple été déterminé expérimentalement que la dérive en fonction de la température est proche d'un élément de la barrette DTC par degré.

Un autre point qui demande vérification est la stabilité à long terme de l'appareil. Nous n'avons aucune idée précise à priori de l'ordre de grandeur des dérives possibles et de leur échelle de temps pour des durées dépassant quelques jours: ce genre de mesures requiert de longues périodes sans intervention dans un environnement contrôlé, deux conditions impossible à réaliser au laboratoire. Comme il était par ailleurs nécessaire d'assurer une bonne précision aux utilisateurs du télescope, il fut décidé d'étalonner l'échelle en fréquence du spectrographe une fois par semaine, en principe pendant la maintenance. Notons qu'ici le terme étalonnage est parfaitement approprié étant donné la précision atteinte.

Cette mesure est effectuée de la façon suivante:

1. Cinq raies de fréquences ν_j , $j = 1..5$ connues sont injectées dans le spectromètre à

l'aide d'un générateur de laboratoire. Pour chacun des spectres obtenus, notons $V_i(\nu_j)$ la tension mesurée sur le canal i .

2. La radiofréquence est coupée et un spectre $V_i(Z)$ est acquis.

3. Une source de bruit est connectée et on acquiert un spectre $V_i(B)$.

Ensuite chaque $V_i(\nu_j)$ est remplacé par $(V_i(\nu_j) - V_i(Z))/(V_i(B) - V_i(Z))$. Le but de cette opération est d'éliminer le courant d'obscurité et la lumière parasite par soustraction de $V_i(Z)$, et de compenser la non-uniformité des gains des éléments par la division. On utilise dans ce dernier cas le fait que la variation de densité de puissance de la source de bruit sur quelques MHz est très inférieure aux différences de gain entre deux points voisins du DTC.

Chacun des spectres $V_i(\nu_j)$ est ensuite assimilé à une raie gaussienne dont les paramètres sont ajustés par un programme de minimisation. L'algorithme utilisé a besoin de valeurs initiales raisonnables des paramètres, qui sont la position, l'intensité et la largeur à mi-hauteur. Les deux premiers sont choisis par une simple recherche du maximum du spectre et le dernier est une constante en théorie égale à la résolution du spectromètre. Dans ces conditions, la procédure de minimisation converge rapidement.

Finalement une parabole est ajustée aux valeurs des positions centrales des cinq raies. Les 3 paramètres c_{ref} , $\Delta\nu$ et a_2 de cette parabole correspondent à la formule suivante:

$$\nu(c_i) - \nu_{\text{ref}} = \Delta\nu(c_i - c_{\text{ref}}) + a_2(c_i - c_{\text{ref}})^2$$

dans cette expression $\nu(c_i)$ représente la fréquence centrale du canal i , ν_{ref} la fréquence intermédiaire de référence (342 MHz), Cette paramétrisation donne directement la valeur du canal de référence (c_{ref}) et l'écart en fréquence entre 2 canaux ($\Delta\nu$) utilisés par le logiciel CLASS d'analyse et de réduction de donnée. La valeur de a_2 n'est pas utilisée pour le moment, il en résulte que l'erreur est minimale au centre et nettement plus importante sur les bords, où elle peut atteindre 1.3 Mhz.

Les variables c_{ref} et $\Delta\nu$ sont enregistrées dans les fichiers de données brutes de façon à pouvoir répéter si nécessaire le processus d'étalonnage et de réduction sans avoir besoin de rechercher leur valeur au moment de l'observation. Le programme qui effectue l'étalonnage écrit de plus, avec une estimation de l'erreur, les paramètres trouvés pour chaque gaussienne et pour la parabole dans un fichier d'archives, pour permettre une analyse ultérieure des dérives de l'appareil.

4.5.7 Dérives à long terme du spectrographe

La période d'observation du spectrographe dépasse maintenant 12 mois, ce qui permet d'effectuer une analyse relativement fine de l'évolution à long terme de l'appareil. Seules sont considérées les mesures effectuées à partir du 15 Août 1990, à cette date la version définitive du programme d'étalonnage est entrée en service et les données ont été systématiquement archivées. Les données antérieures au 1^{er} Août sont de toute façon moins fiables, puisque l'oscillateur local du processeur de fréquence intermédiaire n'était pas encore verrouillé sur l'horloge de référence du télescope.

Des trois paramètres c_{ref} , $\Delta\nu$ et a_2 définis au paragraphe précédent, seul le premier varie de façon mesurable, supérieure à 2 fois l'incertitude estimée, avec le temps. On peut considérer que $\Delta\nu$ et a_2 sont des constantes valant respectivement $-584490 \pm 20 \text{ Hz.canal}^{-1}$ et $6.54 \pm 0.08 \text{ Hz.canal}^{-2}$ et qu'elles ne sont susceptibles de varier qu'en cas d'intervention

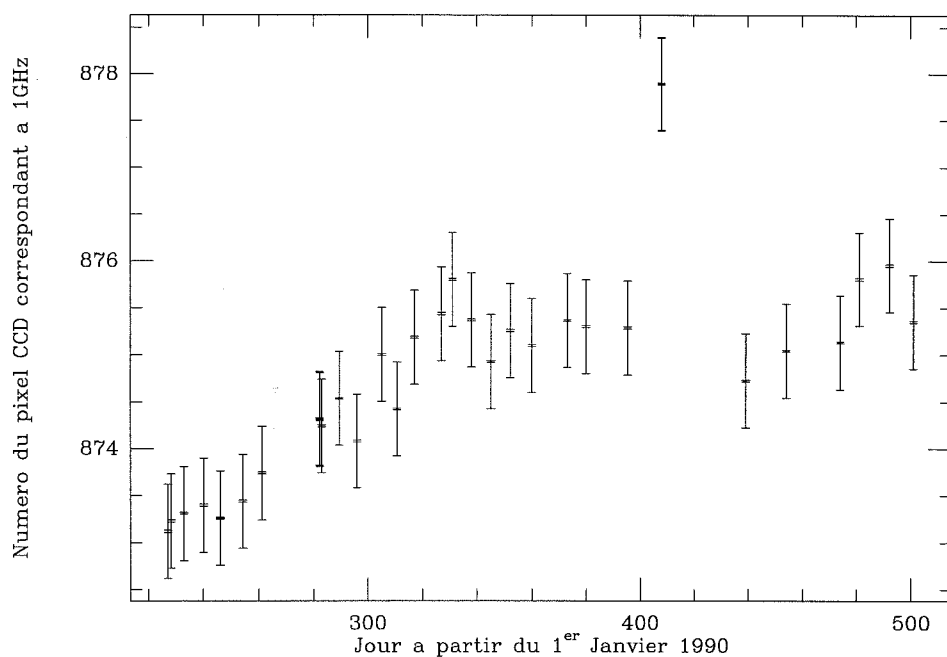


Figure 4.13: Numéro de l'élément du DTC correspondant à une fréquence acoustique de 1 GHz en fonction du temps. Les petites barres d'erreur correspondent à l'incertitude estimée par le programme d'étalonnage, les grandes à un écart de température de ± 0.5 K.

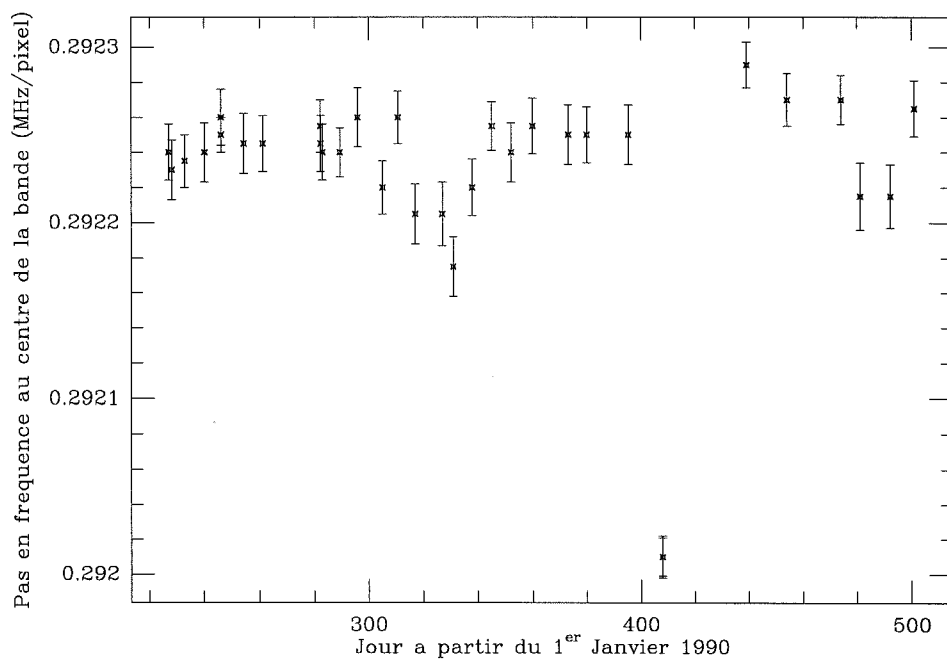


Figure 4.14: Pas en fréquence du spectrographe en fonction du temps, mesuré au centre de la bande.

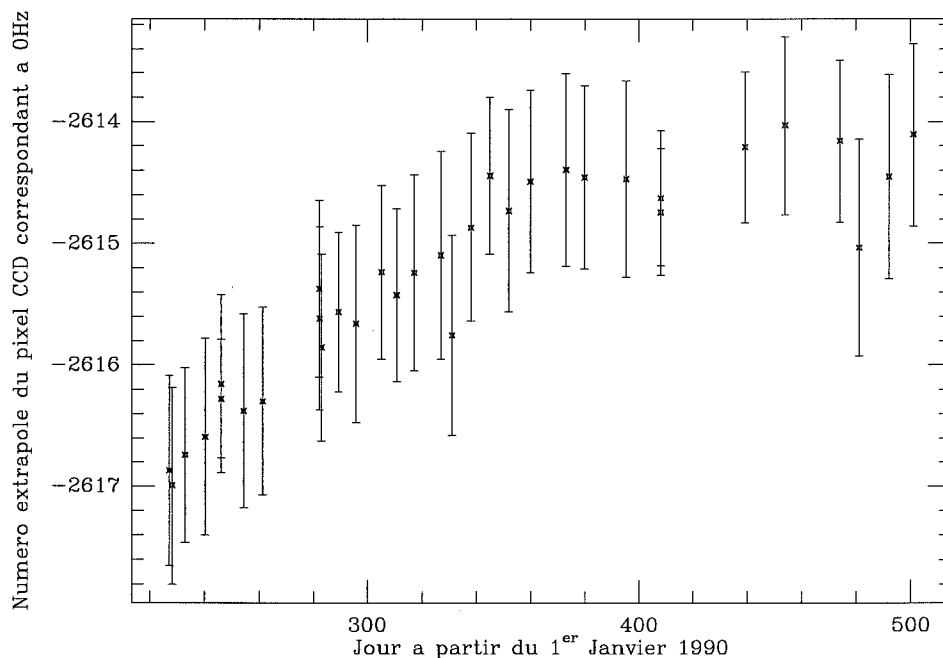


Figure 4.15: Numéro extrapolé de l'élément du DTC correspondant à une fréquence acoustique nulle.

sur la partie optique. La valeur de $\Delta\nu$ est négative à cause du retournement de spectre effectué par le processeur de fréquence intermédiaire.

Le cas de c_{ref} est différent, comme l'illustre la figure 4.13. Le spectrographe étant placé dans une salle climatisée dont la température varie tout au plus de ± 1 K, les phénomènes thermiques ne peuvent à eux seuls expliquer la variation observée. On peut aussi exclure nombre de causes, comme une augmentation de la longueur d'onde laser, par le simple fait que c_{ref} et $\Delta\nu$ varieraient dans les mêmes proportions. Les causes de cette dérive, qui correspond à une diminution de la fréquence centrale du spectrographe, sont par conséquent mécaniques; elles correspondent à un déplacement d'environ $30 \mu\text{m}$ en 6 mois de la barrette DTC dans le plan focal de la lentille de Fourier, mais cela semble être plutôt la conséquence d'un léger mouvement de rotation du miroir qui replie le faisceau.

L'évolution ultérieure de l'appareil n'est pas montrée ici parce qu'un problème avec le système de régulation de température de la diode laser a requis une intervention sur la partie optique et que la bande a été recentrée à cette occasion. Recoller les deux jeux de données n'aurait donc aucun sens; indiquons seulement que la dérive de la fréquence centrale, après avoir continué sur le même rythme pendant environ six mois, a diminué au point de devenir non mesurable sur des périodes de moins de deux mois.

4.6 Conclusion

La comparaison avec les batteries de filtre montre bien que le spectrographe acousto-optique est capable de donner des résultats de la même qualité. Comme le montrent les tracés de variance d'Allan, il atteint le même niveau de stabilité.

Cependant, et bien que ce soit un problème mineur en pratique, la forme du filtre équivalent ne devrait pas exhiber d'ailes larges à bas niveau pour être complètement satisfaisante. Mais la limitation sur laquelle on risque de tomber le plus fréquemment en pratique est celle de la dynamique. Les gains dans la partie radiofréquence ont du être ajustés avec le plus grand soin pour que le photodétecteur travaille dans la gamme appropriée et la marge de sécurité est faible.

Cela ne doit pas nous faire faire oublier que le spectrographe est utilisé couramment par les observateurs qui viennent au télescope et qu'il a démontré une bonne fiabilité [19]. En parallèle avec les deux batteries de filtres, il permet d'observer en large bande simultanément avec trois récepteurs. C'est une possibilité unique qu'offre le radiotélescope de 30 mètres et qui est particulièrement appréciée par les astronomes étudiant les galaxies extérieures.

Chapitre 5

Système solaire

5.1 Introduction

L'étude du système solaire et la radioastronomie millimétrique ont longtemps été des domaines pratiquement disjoints; dans cette gamme de longueur d'onde, les astronomes ont en effet surtout regardé les planètes comme des sources ponctuelles parmi les plus fortes du ciel et les ont utilisées pour des vérifications techniques, comme le pointage et la mesure du coefficient de couplage du lobe principal. Il y avait de bonnes raisons à cela, une antenne de 12 mètres comme celle de Kitt Peak possède un lobe principal supérieur à la minute d'arc à 3 mm, et les récepteurs étaient extrêmement bruyants, donnant des températures système un ordre de grandeur supérieures à celles que l'on atteint aujourd'hui. Dans ces conditions, on ne pouvait observer que des sources étendues et qui seraient de nos jours considérées comme très fortes; dans le système solaire, cela s'était réduit à peu de choses à part les observations de la comète Kohoutek en 1973 et les détections alors annoncées sont toujours considérées comme douteuses.

Comme nous allons le voir, et cela pourrait servir de conclusion à ce chapitre, cette situation a bien changé; les observations de Titan [3], d'Io [4], de Mars [5], de comètes [6], et, à la surprise de tous ceux qui avaient déclarées vaines les tentatives d'observer les planètes géantes, de Neptune [23] constituent des succès récents et importants. Il n'est pas exagéré de dire que le radiotélescope de 30 mètres de l'IRAM a joué un rôle de tout premier plan dans cette évolution; elle est le résultat de la combinaison de l'augmentation de la surface collectrice avec la réduction du bruit des récepteurs qui a fait gagner entre 3 et 4 ordres de grandeur sur le temps d'observation nécessaire pour obtenir le même résultat.

5.2 Titan

5.2.1 Introduction

Titan, satellite de Saturne, fut découvert par Huyghens en 1655. Nos connaissances sur cet objet se réduisaient essentiellement à ses paramètres orbitaux et au fait que c'est l'un des plus volumineux satellites du système solaire jusqu'en 1944, lorsque G.Kuiper mit en évidence la présence de méthane gazeux et donc d'une atmosphère autour de Titan.

L'essentiel de nos connaissances sur Titan vient à l'heure actuelle de la sonde Voyager 1, qui s'approcha à environ 4000 km de sa surface en Novembre 1980. Ce survol fut très riche en enseignements, il prouva tout d'abord que Titan est, avec la Terre, le seul corps du

système solaire pourvu d'une atmosphère dense majoritairement composée d'azote et que la pression au sol est proche de 1.5 bar. Par contre, ce fut aussi une déception dans la mesure où la présence de couches nuageuses opaques de couleur rouge-brun ne permit pas d'en observer la surface, l'un des objectifs importants de cette mission. La composition de ces aérosols atmosphériques n'est pas connue mais des expériences de laboratoire [38] suggèrent qu'il s'agit de composés organiques complexes. En fait, parmi les satellites observés par les sondes Voyager, c'est celui dont on sait le moins de choses sur le plan géologique puisqu'on ne connaît même pas l'état de sa surface avec certitude. Les conditions de température et de pression observées sont tout à fait compatibles avec la présence d'un océan dont le constituant principal serait un mélange d'éthane C_2H_6 et de méthane CH_4 [39, 40]. Les mesures d'écho radar effectuées en 1989 [41] favorisent l'hypothèse d'une surface solide mais ne permettent en aucun cas d'exclure qu'elle soit au moins partiellement liquide.

Les spectres infrarouges pris par la sonde avaient aussi révélé la présence dans la stratosphère de Titan d'hydrocarbures et de composés formés à partir de l'azote et du méthane: C_3H_8 , C_3H_4 , C_2N_2 , HCN, HC_3N , etc.

Enfin, d'autres observations en infrarouge depuis la Terre avaient permis la détection du monoxyde de carbone CO [42].

Dans ces conditions, il paraissait naturel d'essayer d'observer les mêmes composés dans des longueurs d'onde différentes depuis la Terre et en particulier en radioastronomie millimétrique. Ces observations présentent en effet un certain nombre d'avantages: d'une part la présence dans l'atmosphère de Titan de nuages et d'aérosols qui peuvent diffuser le rayonnement à courte longueur d'onde rend la modélisation du transfert de rayonnement plus sujette à caution, d'autre part la résolution spectrale de ces observations était très limitée et ne permettait pas d'observer les raies dans la stratosphère, région intéressante du point de vue photochimique et pour laquelle les théoriciens ont développé des modèles sophistiqués [43, 44]. Les observations infrarouge de Voyager n'avaient permis d'observer la stratosphère que par quelques dizaines de spectres pris en visée au limbe, celles effectuées depuis la Terre sont quasiment limitées au corps qui ne sont pas détruits avant de diffuser vers les couches basses de l'atmosphère, où ils condensent presque tous.

Les observations millimétriques n'en sont pas pour autant très faciles, la faible taille de Titan requiert des antennes très grandes pour limiter le facteur de dilution dans le lobe et des récepteurs peu bruyants. Les modélisations du transfert montraient que l'on pouvait s'attendre à observer dont la largeur à la base varie de quelques dizaines de mégahertz à quelques gigahertz (auquel cas on ne peut pas observer la raie dans son entier), autour d'un pic central étroit et relativement intense [1]. L'information sur la distribution verticale d'un corps est donnée par le profil de la raie, qu'il faut donc déterminer avec précision ce qui explique le temps d'intégration important qu'il faut y consacrer.

La proximité de Saturne, environ 400 fois plus brillant, ne s'est jamais révélée être un problème sérieux pour les observations spectrales. Par contre le fait qu'il soit passé en 1989 par le point le plus bas de son orbite vu depuis l'hémisphère Nord, culminant vers 30 degrés depuis le Pico Veleta, est défavorable: le nombre d'heures d'observation quotidiennes est limité et l'absorption atmosphérique dégrade la sensibilité, surtout à haute fréquence. La distance de la Terre à Saturne peut varier entre 9 unités astronomiques à l'opposition et 11 à la conjonction, il faut en tenir compte car, toutes choses égales par ailleurs, le temps d'observation nécessaire varie comme la quatrième puissance de cette distance; l'opposition, dont la date retarde chaque années de 12 jours, est malheureusement tombée pendant l'été ces dernières années, quand l'absorption atmosphérique est la plus forte, et nous avons été obligés de choisir des compromis sur les dates d'observation. Enfin, vu depuis la Terre,

Saturne s'est trouvé dans le centre galactique, région riche en nuages moléculaires qui ont été, si l'on peut dire, la source de désagréments lors des observations du monoxyde de carbone.

5.2.2 Transfert de rayonnement

Les raies que l'on observe sont dûes non seulement au contraste de température entre la stratosphère et la surface, mais aussi au fait que l'altitude jusqu'à laquelle l'atmosphère peut être optiquement épaisse est loin d'être négligeable face au rayon de Titan. Cet effet peut augmenter le contraste des raies d'environ 30 % pour la première transition de l'acide cyanhydrique à 88.6 GHz et pour en tenir compte, il faut effectuer le calcul de transfert de rayonnement en géométrie sphérique et non sous une incidence moyenne en géométrie plan-parallèle [1], ce qui rend les calculs nettement plus lourds.

Le temps de calcul nécessaire pour intégrer l'équation de transfert reste toutefois très raisonnable dans le domaine millimétrique: une dizaine de secondes sur un ordinateur personnel rapide de 1991. Par contre, la version infrarouge du programme, qui peut avoir à prendre en compte des milliers de raies et une grille en fréquence de quelques centaines de milliers de points, exige des moyens informatiques nettement plus lourds, tant en puissance de calcul qu'en espace disque.

5.2.3 Observations

Trois molécules ont été observées jusqu'ici sur Titan dans le domaine millimétrique: ce sont l'acide cyanhydrique HCN, le monoxyde de carbone CO et récemment le cyanoacétylène HC_3N [9]. Nous ne nous étendrons pas sur ce dernier, les deux raies observées vers 145 et 218 GHz sont en effet plus étroites que prévu puisqu'elles n'occupent qu'un seul canal de 1 MHz de large; cela rend vaine toute tentative d'analyse mais indique probablement l'existence d'un fort gradient de concentration dans la stratosphère et rend absolument indispensable l'utilisation d'un modèle de transfert sphérique si on arrive à extraire un profil de raies à partir d'observations à plus haute résolution.

En règle générale, les observations ne permettent pas de mesurer le flux absolu de Titan, il y a donc une incertitude sur le niveau moyen du signal. C'est évident de tous les spectres obtenus jusqu'à la mise en service du secondaire oscillant, et même ceux acquis après exhibent une trop forte dispersion du niveau moyen pour pouvoir leur faire la moindre confiance. Cela peut être dû à des instabilités du récepteur ou de l'atmosphère qui ne sont pas bien rejetés par la fréquence maximale de commutation du secondaire en mode spectral: 0.25 Hz. Dans ces conditions, on rajoute aux spectres observés la constante nécessaire pour que leur valeur moyenne sur la bande soit égale à celle du spectre synthétique. Ce n'est qu'ensuite que l'on essaie de minimiser la différence; cela signifie en pratique que l'on ne peut tirer de l'information que des ailes de raie où la pente est suffisamment forte pour ne pas être dominée par le bruit.

Il est clair d'après les spectres de Voyager que l'abondance de certains composés dépend de la latitude; il est bien évidemment impossible de faire la distinction entre la contribution des pôles et de l'équateur dans les observations effectuées depuis la Terre. On n'en déduira donc que des valeurs moyennées sur le disque, ce qui privilégie les régions équatoriales même si le transfert en géométrie sphérique atténue quelque peu cet effet de sélection.

5.2.4 Acide cyanhydrique

La première transition de l'acide cyanhydrique à 88.63 GHz fut détectée en Septembre 1986 à l'IRAM. Cette détection, analysée grossièrement pour une présentation quelques semaines plus tard [7], et malgré des conditions d'observation exécrables, indiquait déjà que le rapport de mélange n'était pas constant avec l'altitude.

Les observations effectuées en Mai 1987 ont confirmé, en la précisant grâce à une meilleure détermination du profil de la raie, cette première conclusion. Ils indiquent que le rapport de mélange de HCN croît avec l'altitude et donnent des contraintes sur le coefficient de diffusion turbulente dans l'atmosphère [3].

Des observations ultérieures de la troisième transition à 266.9 GHz ont une nouvelle fois vérifié ces résultats mais, à cause de problèmes de modélisation quand la raie est plus large que la bande passante, elles n'ont pas permis de les améliorer [45].

Il se trouve malheureusement que la deuxième transition, vers 177.25 GHz, n'est pas accessible depuis la Terre. Une raie intense de la vapeur d'eau située à 183 GHz rend en effet l'atmosphère totalement opaque la majeure partie du temps et l'étalonnage très incertain dans les rares occasions où l'on pourrait observer.

5.2.5 Monoxyde de carbone

La première observation du monoxyde de carbone sur Titan date de Mai 1987, en parallèle avec la deuxième observation de l'acide cyanhydrique. Le rapport signal sur bruit très limité ne permet pas d'en tirer de conclusions définitives [2] mais apparaît en contradiction totale avec la valeur infrarouge [42], nettement plus forte.

Pour améliorer ce résultat, les trois premières transitions du monoxyde de carbone, de fréquences respectives 115.27, 230.54 et 345.8 GHz ont été observées en Avril 1991 avec le radiotélescope de 30 m. A cause d'un rapport signal sur bruit nettement plus médiocre et de plus fortes incertitudes d'étalonnage sur la troisième, nous ne considérerons ici que les 2 premières. La première de ces raies est toujours bien plus faible que ce que la détermination infrarouge laissait prévoir, mais nettement plus forte qu'en Mai 1987; il faut supposer que nous avons lors de cette première observation été victimes d'un problème de rapport de gain entre les deux bandes, ce n'est pas surprenant puisque nous utilisions alors un récepteur SIS récemment installé et qu'à cette fréquence il fallait alors utiliser la bande supérieure dont on ne savait pas encore qu'elle pouvait être défavorisée par rapport à la bande inférieure. Ces problèmes sont aujourd'hui maîtrisés.

Revenant donc aux données obtenues en 1991, nous les avons analysées avec les mêmes méthodes que celles utilisées pour l'acide cyanhydrique [3] dans le but d'en déduire la distribution du monoxyde de carbone dans la stratosphère. Les premiers résultats présentés ici sont préliminaires dans la mesure où les barres d'erreur n'ont pas encore été déterminées.

Comme dans le cas de l'acide cyanhydrique, le spectre obtenu est incompatible avec une distribution verticale uniforme, mais cette fois le rapport de mélange décroît avec l'altitude dans la stratosphère. Comme les observations millimétriques ne sondent pas les couches profondes de l'atmosphère, nous avons fixé le rapport de mélange aux pressions supérieures à 10 mb à la valeur $r_0 = 6 \cdot 10^{-5}$ donnée par les mesures infrarouge [42], il devrait être possible de remonter ce seuil à 100 mb, mais cette gamme de pression supplémentaire n'affecte que les bords du spectre où la pente est faible et entachée d'une incertitude de mesure excessive. Pour les pressions inférieures, nous avons supposé ensuite que le rapport de mélange r variait comme une loi de puissance de la pression, $r = r_0(P/P_0)^\alpha$. Enfin, nous avons décidé de

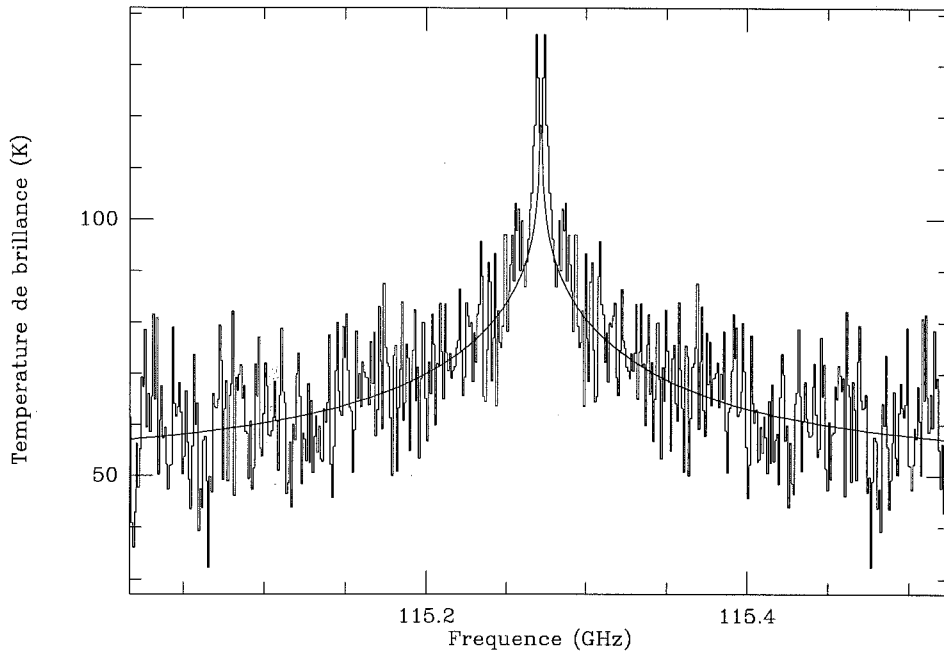


Figure 5.1: Spectre de Titan dans la première transition de CO

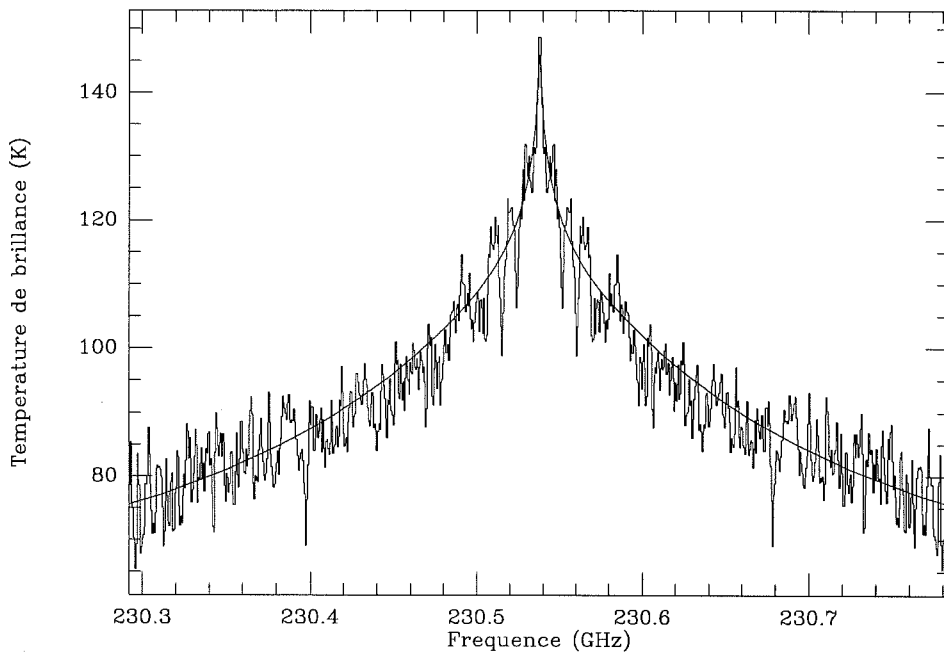


Figure 5.2: Spectre de Titan dans la deuxième transition de CO

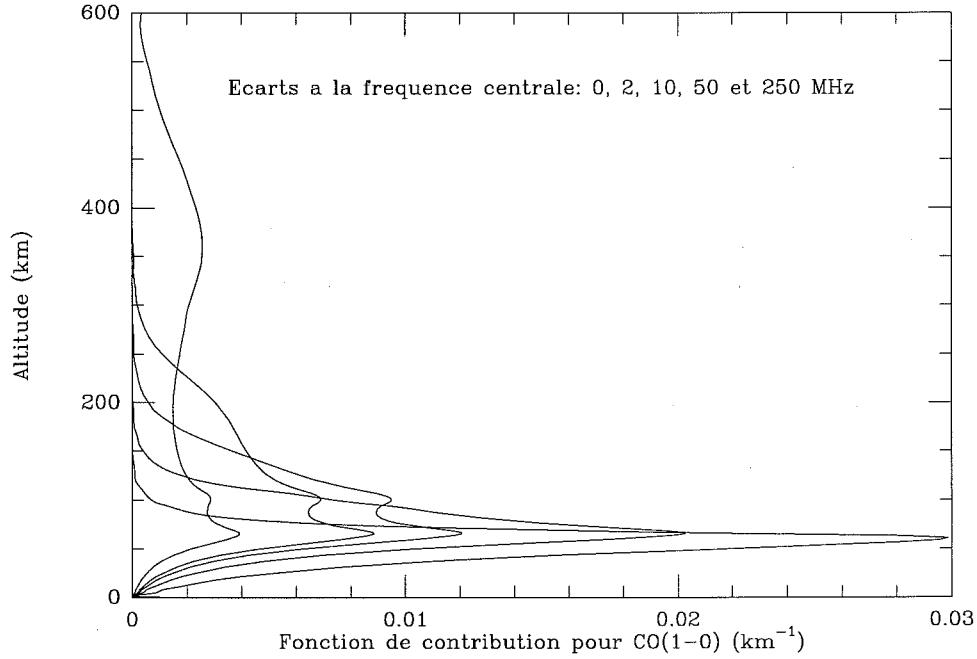


Figure 5.3: Fonctions de contribution dans la première transition de CO

couper cette loi de puissance vers 0.1 mb puisque la largeur de raie correspondante devenait plus faible que la résolution, et que le rapport signal sur bruit limité ne nous permettait pas de nous aventurer plus loin. Cette gamme de pression correspond à des altitudes comprises entre 100 et 300 km.

Le modèle n'a donc plus qu'un seul paramètre libre, α , ce qui est justifié pour une première approche mais devra être amélioré par la suite. Le meilleur accord avec les observations, montré par les figures 5.1 et 5.2 s'obtient pour une concentration de monoxyde de carbone égale à $2 \cdot 10^{-5}$ dans la limite des basses pressions, soit $\alpha \simeq 0.24$. Il est visible sur les figures que les spectres ont été repliés autour de la fréquence centrale pour améliorer le rapport signal sur bruit, cette opération n'affecte en rien le résultat [46]. Malheureusement un pic de bruit près du centre de la première raie de CO, peut-être dû à une interférence avec le récepteur submillimétrique, limite un peu son utilité. Ces canaux ont été ignorés lors de la détermination du profil, ils ont été laissés sur la figure uniquement pour donner un exemple des problèmes auxquels on peut se trouver confrontés.

Les fonctions de contribution correspondant aux distributions obtenues sont montrées dans les figures 5.3 et 5.4. Elles ont été calculées sans tenir compte de la résolution du spectrographe pour donner une idée de la gamme sondée au centre de la raie si l'on avait un bon rapport signal sur bruit à haute résolution. Elles montrent que cette première détermination des paramètres est raisonnable car elle dépend surtout des écarts en fréquence compris entre 2 et 50 MHz, qui correspondent aux altitudes comprises entre 100 et 300 km. On devrait pouvoir agrandir cette zone dans les deux directions avec un meilleur rapport signal sur bruit.

Si on essaie de calculer le flux de CO en particules par centimètre carrés déduit du gradient de concentration déterminé ici, on obtient environ $10^{10} \text{ cm}^{-2} \cdot \text{s}^{-1}$. Ce flux est énorme et il impliquerait une disparition totale du monoxyde de carbone en environ 50000 ans, 5

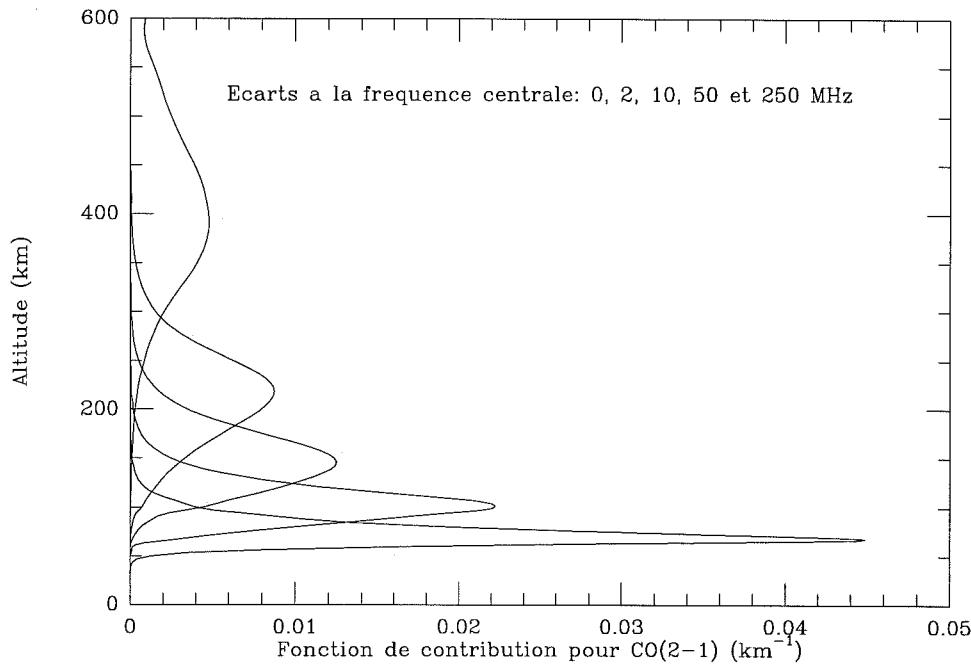


Figure 5.4: Fonctions de contribution dans la deuxième transition de CO

ordres de grandeur de moins que l'âge de Titan. Cette différence est telle qu'aucune erreur de mesure ne saurait l'expliquer.

La première explication qui vient à l'esprit est de supposer qu'il existe une source de CO au niveau de la surface qui fournit ce flux. Cette hypothèse est cependant difficile à soutenir car la quantité de CO que l'on peut dissoudre dans l'océan n'est qu'un ordre de grandeur supérieure à celle présente dans l'atmosphère, et de plus, on l'a déjà dit, des expériences d'écho radar indiquent que la surface est en majeure partie solide.

L'autre hypothèse que l'on peut faire est qu'il existe un mécanisme de recyclage du monoxyde de carbone. De tels mécanismes ont été proposés [47] pour la basse stratosphère dans lesquels les particules et aérosols fixent les molécules de CO à une altitude et les libèrent à plus basse altitude. Ce type de mécanisme semble le plus probable étant donné qu'on observe des brouillards organiques jusqu'à des altitudes de plusieurs centaines de kilomètres, c'est à dire dans la gamme dans laquelle nous avons déterminé que le gradient de concentration doit exister.

Cette analyse est encore très imparfaite, il n'en reste pas moins vrai que la conclusion à laquelle on arrive, que la concentration de monoxyde de carbone ne peut pas être constante avec l'altitude, est assurée après avoir essayé des distributions verticales constantes dans toute la gamme d'erreur de la détermination infrarouge et même en dehors: ce n'est pas un simple facteur d'échelle que l'on pourrait expliquer par une erreur d'étalonnage mais le profil du spectre qui est radicalement différent et qu'il est impossible d'ajuster sur les deux raies simultanément. Pour améliorer de façon sensible ce résultat, il faudrait obtenir un meilleur rapport signal sur bruit et une observation à haute résolution du profil du pic central. Une telle observation demanderait cependant pour le moment plusieurs dizaines d'heures d'observation, ce qui peut être considéré comme excessif.

5.3 Autres observations

La suite de ce chapitre doit avant tout être regardée comme une introduction aux articles joints à ce mémoire; on y trouvera certaines des observations d'objets du système solaire auxquelles j'ai directement participé. Quand le projet était trop court pour justifier la venue d'un collaborateur, j'ai effectué moi-même les observations; j'y ai sinon participé de façon d'autant plus active que ces observations représentaient souvent le premier contact de l'observateur avec le radiotélescope, quand ce n'était pas avec la radioastronomie tout court. J'ai aussi pris part au moins aux premières étapes du dépouillement et de l'analyse des données. Par contre, au contraire de Titan pour lequel j'ai écrit le programme de transfert de rayonnement spécifique qui est à la base de tous les calculs de modélisation et de comparaison de spectres, je n'ai pas participé à l'interprétation de ces données. Je ne prétends pas être un spécialiste de ces autres corps et l'analyse des phénomènes physiques sous-jacents requiert souvent des programmes de modélisation spécifiques dont je ne dispose pas.

En classant les objets par ordre décroissant de distance du périhélie, nous allons présenter brièvement les observations et les principaux résultats obtenus sur quelques comètes, Io, Mars, et finalement quelques comètes.

5.3.1 Io

On sait depuis les années 70 qu'Io possède une atmosphère ténue, grâce à la détection d'une ionosphère par la sonde Pioneer 10. D'autres observations avaient permis de conclure à la présence de dioxyde de soufre, d'une part sous forme solide à la surface [48], d'autre part sous forme gazeuse près des volcans grâce aux sondes Voyager.

L'observation directe [49] avec le radiotélescope de 30 mètres du dioxyde de soufre en phase gazeuse a constitué la première détection globale de l'atmosphère neutre d'Io.

Dans le cas d'Io, la réduction des données demande un soin particulier parce que le logiciel de poursuite des satellites de l'antenne ne tient pas compte de la vitesse relative de la source par rapport à Jupiter. La raison de ce problème est que les éphémérides que nous utilisons, celles publiées régulièrement par le Bureau des Longitudes de Paris, ne permettent pas de calculer de façon simple la vitesse radiale d'un satellite; il semble bien que ce soit le cas de toutes les sources d'éphémérides disponibles à l'heure actuelle.

Pour Titan, dont la vitesse sur son orbite ne dépasse pas 6 km.s^{-1} et la période de révolution dure 16 jours, on peut se contenter de corriger la vitesse la somme des spectres après chaque séance d'observations. Comme on l'a vu, à l'exception peut-être du pic central que nous n'avons pour le moment jamais réussi à mesurer avec un rapport signal sur bruit permettant de l'analyser, le fait que les raies soient larges et que l'on utilise les spectrographes à faible résolution permet de simplifier considérablement le problème.

Le cas d'Io est complètement différent, on s'attend à des raies étroites dont la largeur est déterminée par la vitesse thermique en raison de la faible pression. Or sa période orbitale n'est que de 42 h et sa vitesse par rapport à Jupiter peut atteindre 17 km.s^{-1} , bien plus que la largeur de raie attendue, et elle varie rapidement. Pour cette, les temps d'intégration des spectres individuels n'ont jamais dépassé 2 minutes; ces spectres ont ensuite été rééchantillonné sur une grille en fréquence commune avant de les additionner. Cette procédure modifie la forme équivalente des filtres et d'autres paramètres des spectres; bien que la plupart des discussions entre les membres de l'équipe aient eu lieu par courrier électronique, j'ai cru nécessaire dans ce cas précis d'écrire un rapport détaillé, poussé par l'absence apparente de références à des cas similaires dans la littérature et le besoin d'obtenir l'accord général

lors du dépouillement des données. Ce rapport, écrit dans les semaines qui ont suivi les observations de Janvier 1991, constitue l'appendice C; il analyse les effets du rééchantillonnage sous des hypothèses réalistes.

Malgré les incertitudes et la complexité de l'analyse [4], ces observations représentent un progrès certain; parmi les modèles proposés jusqu'alors, elles privilégient l'hypothèse d'une atmosphère d'origine essentiellement volcanique.

5.3.2 Mars

Nous avons saisi l'occasion que représentait l'opposition de Mars en Septembre 1988 pour l'observer dans les conditions les plus favorables; à ce moment là, sa taille atteignait en effet presque 24 secondes d'arc, soit le double du lobe du télescope à 230 GHz, ce qui permet une cartographie grossière. Lors de ces observations, il faut apporter une attention particulière au pointage dont la précision est critique.

Le premier objectif était de savoir si la distribution du monoxyde de carbone était uniforme ou non [10] par l'analyse simultanée des raies du monoxyde de carbone ^{12}CO et de son isotope ^{13}CO . L'étude détaillée du premier, sur lequel nous avons obtenu en tout et pour tout seulement 28 spectres d'une minute chacun, a cependant démontré, que même avec une résolution spectrale de 100 kHz, on pouvait déterminer la position du centre de la raie avec une précision de l'ordre de 15 kHz, soit à peu près 20 m.s^{-1} .

La précision sur la mesure des vitesses est telle qu'elle dépasse les limites pour lesquelles le télescope est conçu, à cette fréquence en effet les synthétiseurs qui génèrent l'oscillateur local donnent après multiplication un pas de 24.48 kHz. Une fois les données corrigées des écarts systématiques causés par cette limitation, on peut extraire des spectres une carte de la vitesse en fonction de la longitude et de la latitude [5]. Le faible nombre de points de la carte, 14, et la résolution spatiale limitée qui convolue les prédictions théoriques ne permettent qu'une comparaison des caractéristiques globales de la circulation atmosphérique aux altitudes moyennes sondées; on peut cependant en conclure que ces premières observations confirment les modèles actuels.

5.3.3 Comètes

Un certain nombre de comètes ont été observées ces dernières années au 30 mètres, en commençant par la célèbre comète de Halley au début de sa mise en service. L'absence à ce moment de récepteur à haute fréquence et les problèmes inhérents aux balbutiements d'un instrument de cette importance ne permettent pas de prétendre que l'on ait obtenu de résultats spectaculaires [50].

Suivirent quelques années de sécheresse, au cours desquels les passages des comètes Wilson et Machholz, cette dernière ayant disparu sans laisser de traces lors de son passage au périhélie, ne permirent aucune détection certaine [11].

Après cela, quelques premiers résultats encourageants, avec la détection de HCN et du formaldéhyde H_2CO , furent obtenus sur la comète Brorsen-Metcalf (1989o).

Mais c'est surtout l'année 1990, avec le passage des comètes Austin (1989c1) et Levy (1990c) qui a été l'occasion d'effectuer au 30 mètres des observations intéressantes, en particulier la première détection directe du sulfure d'hydrogène H_2S et du méthanol CH_3OH [6, 14]. Ces deux détections fournissent des contraintes importantes aux astronomes qui s'intéressent à la formation des comètes.

De ces quatre molécules, trois sont considérées comme étant des molécules-mères, c'est à dire qu'elles émanent directement du noyau de la comète et permettent donc de déterminer sa composition chimique. Le cas du formaldéhyde est plus complexe, en raison de la facilité avec laquelle le rayonnement ultraviolet le détruit, il ne devrait être présent que très près du noyau; or son extension spatiale semble très supérieure, ce qui ne peut s'expliquer que s'il existe une source indirecte de formaldéhyde, par exemple s'il est le résultat de la décomposition d'une autre molécule.

Quoiqu'il en soit, la première conséquence de la présence de H_2S est simple, sa très grande volatilité implique que les comètes se sont formées à une température très basse. La limite supérieure de cette température dans un mélange dont le composant principal qui condense est l'eau est encore l'objet de discussions [12].

Quant au méthanol, lui aussi très volatile, le fait qu'il soit détruit par le rayonnement ultraviolet alors qu'il est présent dans le milieu interstellaire en abondance non négligeable indique aussi que les comètes se sont formées dans des régions froides et loin du centre de la nébuleuse primitive, dont elles conserveraient, au moins partiellement, la composition chimique originelle. Une analyse plus détaillée sera publiée ultérieurement.

Ces observations jouent en faveur des théories les plus couramment admises sur la formation des comètes: celles qui disent que ce sont les objets les plus primitifs du système solaire, et qu'elles n'ont pas ou peu évolué depuis leur formation.

Appendice A

Measurements of receiver power stability

A.1 Introduction

During the last technical time session, two hours were allocated during the night between May 30th and 31st to evaluate the stability of the receivers in continuum mode. Such a measurement had been performed on the 3mm Schottky receiver a few months ago.

Because of the failure of the 2mm receiver, only the 3mm and 1.3mm SIS receiver could be checked.

A.2 Technical details

These measurements require a higher sampling frequency than the one normally provided by the standard data acquisition software. This is obtained by putting a clock controller module in the continuum backend crate which interrupts the microprocessor every 2 ms, giving a sampling frequency of exactly 500 Hz (the clock controller is referenced on the station clock). In order to avoid quantization noise problems in the voltage to frequency converters at such a high sampling frequency, it is necessary to have a power level higher than 100 on the monitoring displays.

Special programs are then run, both in the CAMAC microprocessor and on the VAX. Data is dumped onto a disk file for later analysis. In order to permit efficient and simple Fourier analysis, the number of samples recorded is $524288 = 2^{19}$, which is divided into 64 scans of $8192 = 2^{13}$ points.

The power spectrum of each scan is then computed as the square modulus of its FFT and all scans are averaged and normalized in such a way that the resulting spectrum is an estimation of the inverse of the effective noise integrating bandwidth of the receiver, i.e., the factor B in the well-known $1/\sqrt{B\tau}$ expression.

With the values given above, the resolution of the final spectra is 1/16.384 Hz between 0 and 250 Hz (the 0 Hz point is not shown on the graphs however because it lies far outside the scale and has a constant value equal to 16.384 Hz^{-1}).

Three spectra of each receiver were acquired, one on the hot load and the two others on the sky with the antenna stopped at 60° elevation. The local oscillator power of both receivers was optimized before the second on sky scan. A fourth spectrum was acquired

on the distribution box noise source to check acquisition software and verify continuum backend performance.

A.3 Results

The obtained spectra are shown in figures A.1 to A.4. The spectrum obtained on sky in very cloudy conditions with the Schottky receiver is also shown in figure A.5 and noise source spectra appear in figures A.6 and A.7.

All receivers exhibit a strong 1 Hz pump modulation, including a large number of harmonically related lines. They also all show significant mains (50 Hz and mutiple) interference. Some other lines of unknown origin are also present. The first conclusion that can be drawn is that the sensitivity depends strongly on whether the on-off switching procedure effectively rejects or not the pump modulation.

For the 3mm SIS receiver (figs. A.1 and A.2), which was tuned in double sideband with the lower sideband at 94 GHz, it clearly appears that the theoretical bandwidth is never reached and that the noise starts to rise at a very high frequency. The fact that the spectra on the sky and on the hot load are very similar indicates that the receiver limits the sensitivity of the telescope. Experience gained from observation of flare stars (project 141Y90) shows that the performance of this receiver is for now unpredictable, giving a 10 MHz equivalent bandwidth during the first run (two nights) and approximately 200 MHz one week later, at the same frequency and with the same observing procedure. Several unsuccessful attempts to retune the receiver were performed during the first two nights, indicating that, although the problem may be related to tuning, no relationship has been established and the exact cause of the unstabilities has yet to be determined.

The 1.3mm SIS receiver (figs. A.3 and A.4), on the other hand, although it is also affected by 1 and 50 Hz modulation, exhibits a significant difference between the hot load and the sky, and reaches a reasonable fluctuation bandwidth of 400 MHz at 10 Hz. The difference is however noticeable only below 0.5 Hz, indicating that, under good atmospheric conditions, the sensitivity is indeed also limited by the receiver.

The spectra taken on the noise source (figs. A.6 and A.7), although they clearly show that the detectors do not limit the sensitivity of the telescope, reserve some surprises and do not agree with the results obtained just after the installation of the distribution box. There is a quite strong 50 Hz line on the first channel, which apparently disappears when the power supply of the 100 kHz backend is switched off, and sometimes also a line around 40 Hz. The slow rise in the noise level observed towards the higher frequencies is due to the quantization noise of the voltage to frequency converters and is expected.

However the most annoying difference is that previous measurements indicated that the fluctuation bandwidth stayed well above 500 MHz down to below 0.1 Hz. Analysing the data down to lower frequencies with a better spectral resolution yields the spectrum shown in figure A.8, in which two lines appear around 0.12 and 0.24 Hz. Further analysis of the data shows that all continuum channels are very strongly correlated, which implies that the fluctuations are due to a common cause, for now unknown.

A.4 Conclusions

All the preceding spectra show that the sensitivity of the telescope as a whole is far from theoretical values. In order to improve performance, the following points should be ad-

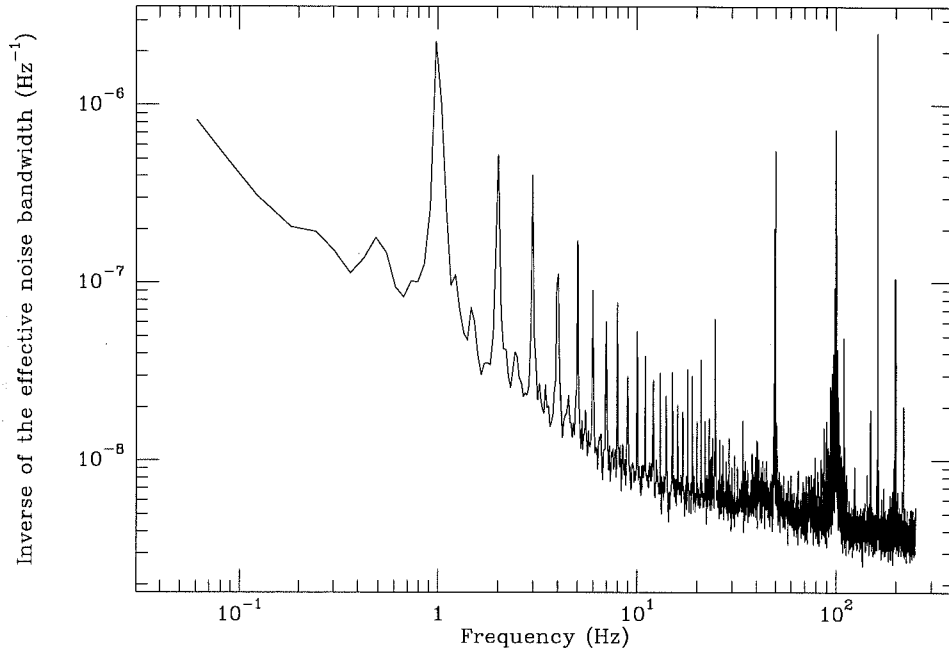


Figure A.1: 3mm SIS receiver on hot load.

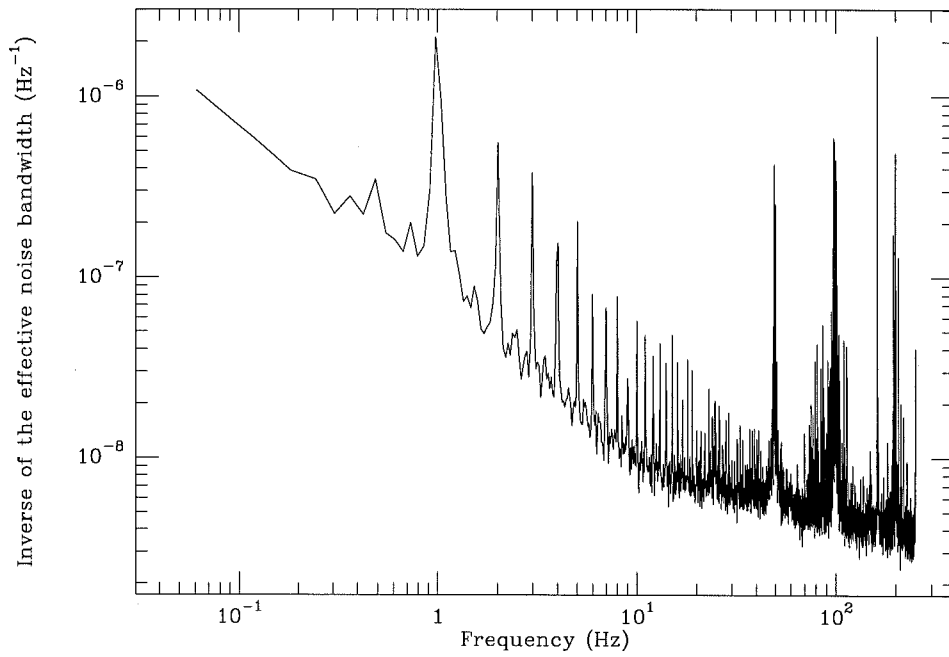


Figure A.2: 3mm SIS receiver on sky.

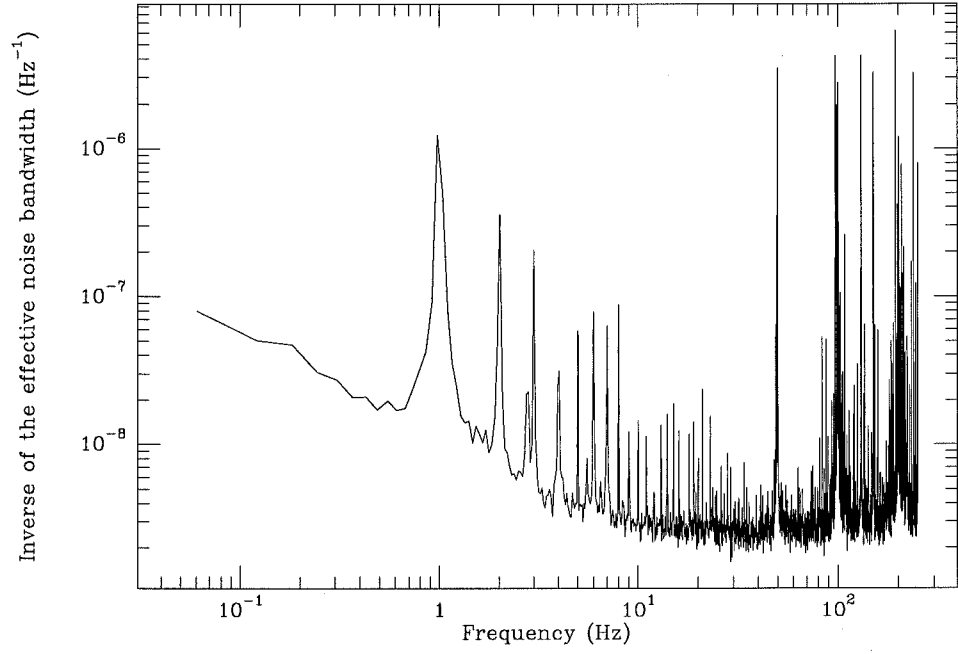


Figure A.3: 1.3mm SIS receiver on hot load.

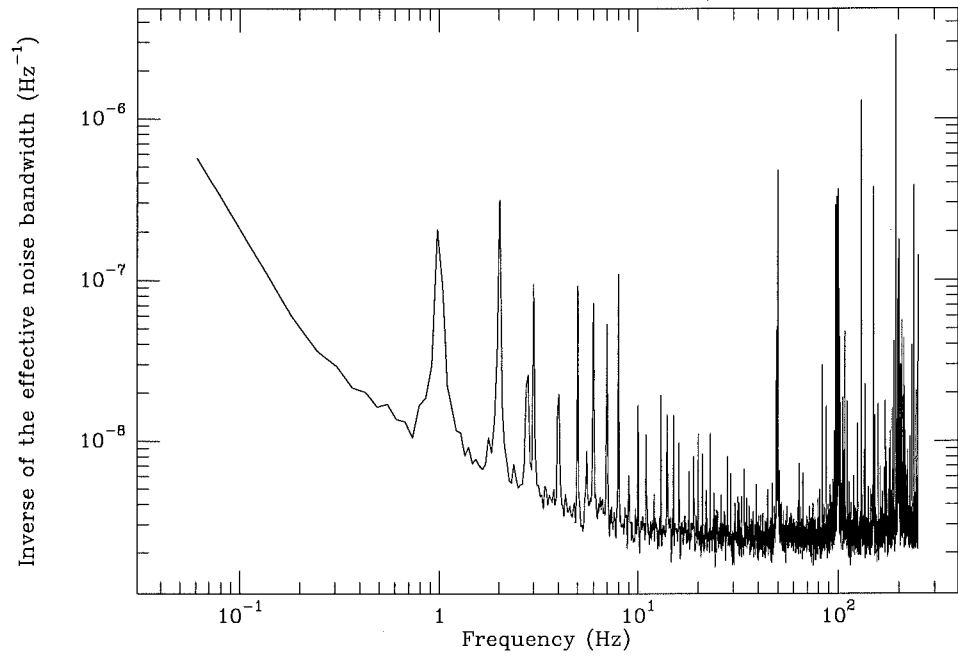


Figure A.4: 1.3mm SIS receiver on sky.

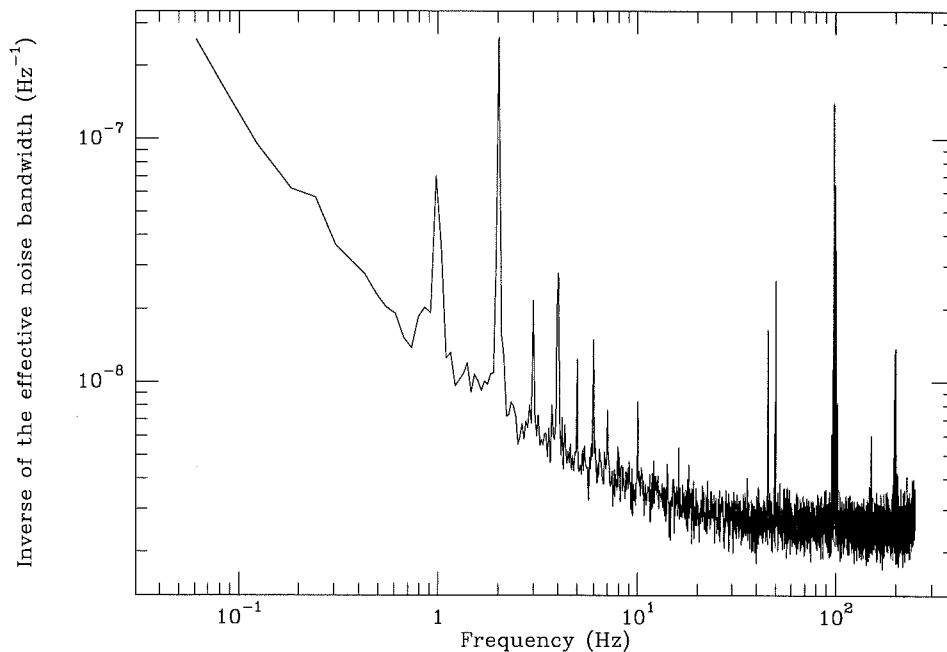


Figure A.5: 3mm Schottky receiver on sky.

dressed:

- understand why the continuum detector performance is different from previous measurements, and try to minimize 50 Hz interference from the power supplies,
- find and remove the cause of the strong mains modulation of the receivers,
- look for the origin of the excessive gain fluctuations of the 3 mm SIS receiver, and try to obtain better and repeatable performance,
- try to minimize the influence of the pump,
- finally develop observing procedures that strongly reject the remaining modulation frequencies, by the application of simple numerical filtering techniques.

Taking into account the fact that the highest realizable switching frequencies are of the order of a few Hertz and assuming that the pump modulation will be the most difficult to minimize, the choice of the switching frequencies is simple: there must be an odd number of phases per pump cycle. This means that the phase time of the wobbler can be chosen for example as 1 s or 200 ms; the smaller value does not necessarily give better results because the proportion of blanking time is much higher: it mostly depends on weather conditions and gain stability of the receiver. The best compromise might turn out to be 3 phases per second.

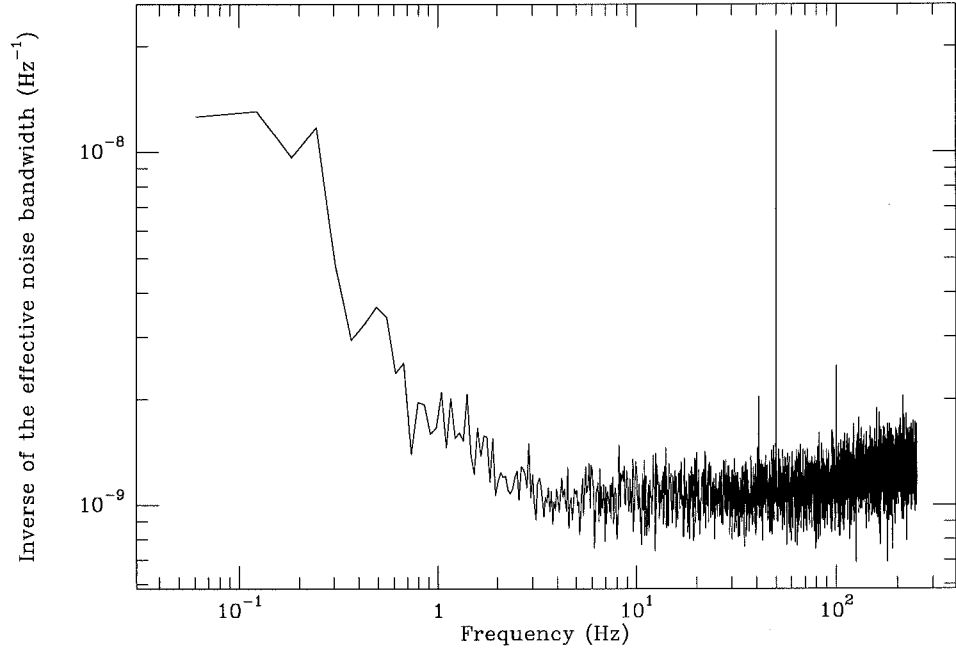


Figure A.6: Noise source on continuum channel 1.

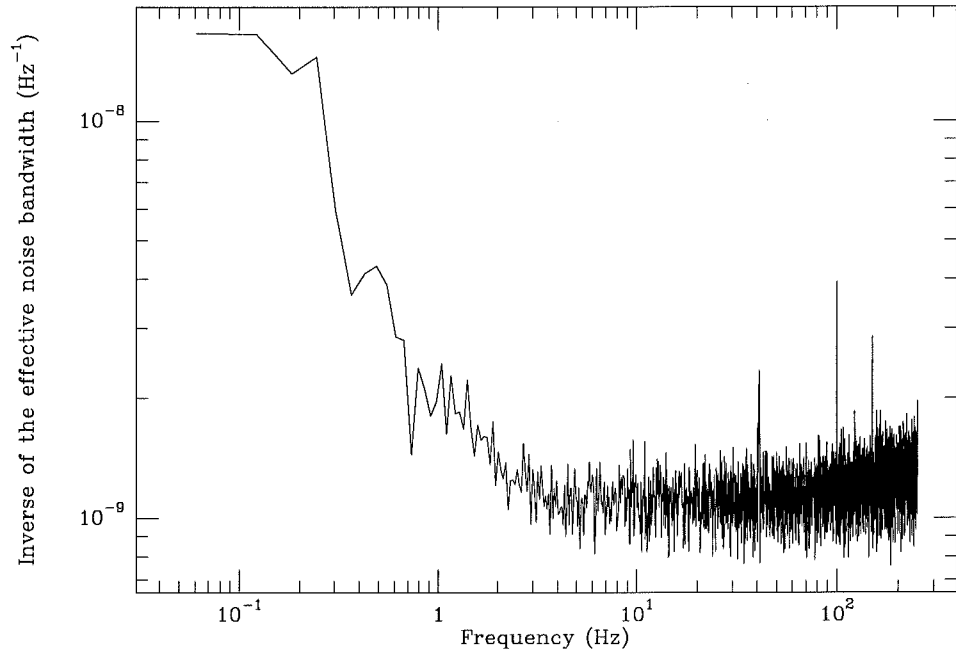


Figure A.7: Noise source on continuum channel 2.

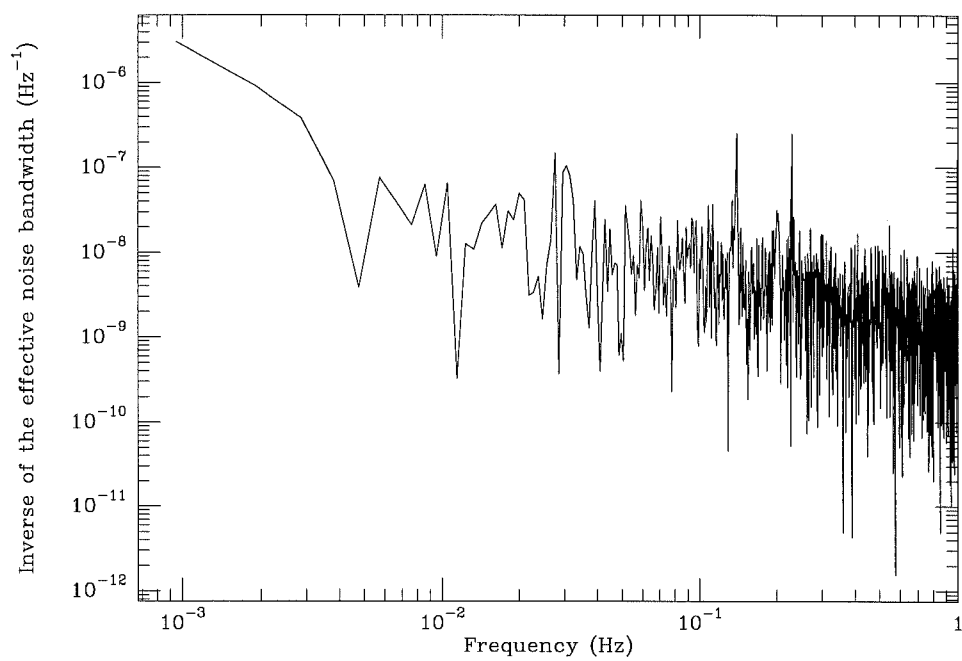


Figure A.8: Noise source on continuum channel 1.

Appendice B

Technical documentation of the AOS acquisition system

Foreword

This report describes the system which was designed to perform the data acquisition and processing for acousto-optic spectrometers at l'Observatoire de Meudon.

This document should be updated on every modification made in order to reflect possible changes and improvements.

Acronyms used in this document

AOS: Acousto-Optic Spectrometer,

ADC: Analog-to-Digital Converter,

CCD: Charge Coupled Device,

SHA: Sample-and-Hold Amplifier,

VFC: Voltage-to-Frequency Converter,

FIFO: First-In First-Out memory (queue),

LCA: Logic Cell Array (user programmable gate array),

CLB: Configurable Logic Block (LCA elements),

IOB: Input-Output Block (LCA elements).

B.1 Overview

At first glance, the system can be divided into two parts: hardware and software. This document only gives clues to the way the software may be written, and gives some advices but is not intended to be a complete documentation of all the programs written for the AOS. It rather concentrates on the processing up to the point where the data of individual CCD pixels reaches microprocessor memory.

The hardware part performs the following tasks:

1. readout of the CCD using the double correlated sampling technique and analog-to-digital conversion,
2. buffering of the data between the ADC and the microprocessor,
3. synchronization with external signals for fast phase switching,
4. synchronization with other AOS (section B.11),
5. monitoring of the laser diode (if the diode control electronics delivers a suitable signal).

Software must set up the hardware to operate as required, and read out the ADC buffer and process the data according to the observation mode.

B.2 Architecture

The electronic is split onto three boards as shown in figure B.1. In this diagram, ovals represent external signals.

The optic board holds only the CCD and its necessary support circuits. Its size is limited by the available space inside the optic box and its power dissipation must be kept low to avoid generating turbulent air flow in front of the CCD, which would obviously cause catastrophic instabilities.

The analog board holds the SHAs and the ADC, as well as some logic for stand-alone operation. It also holds a VFC which was installed with the idea of continuously measuring the power of the laser diode, but which may be used to monitor any quantity represented by a voltage (or a current). This circuit is very well suited to the measure of small, long term variations of a system parameter since it is an integrating converter with excellent differential linearity characteristics and no inherent quantization noise.

The digital board generates all the necessary signals to control the CCD, the SHAs, the ADC and the VFC. It also interfaces with the microprocessor, generates interrupts, controls the integration and holds a FIFO memory to buffer the data between the ADC and the microprocessor. The heart of the system is implemented inside a *Xilinx* LCA [51], which provides a dense and flexible logic architecture for moderately high speed designs, with a much lower power consumption than most competing programmable logic devices.

The CCD and analog boards are linked together through a cable made of a coaxial cable for the amplified CCD output and several wires for power supplies and control signals. The link between the analog and the digital board is realized with a simple 34 wire flat cable, which should be kept as short as possible because its parasitic capacitance combined with the fast edge rates of high-speed CMOS may cause strong current spikes on the power supplies and increase ground noise on the analog board. Five of these connections are for

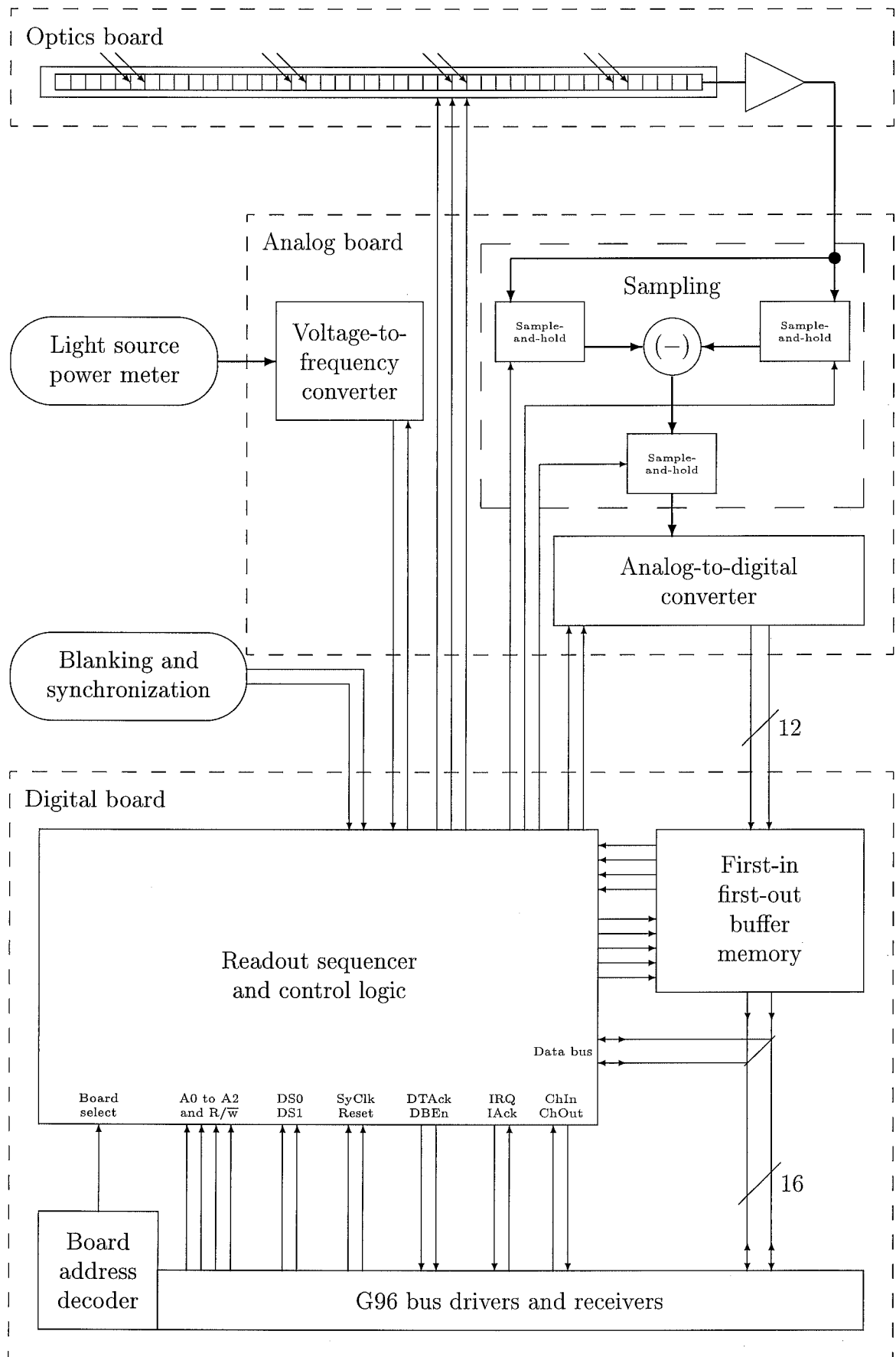


Figure B.1: General diagram of the acquisition system

now not used: one is designated as a spare control signal and the four others allow the use of up to 16 bit ADC.

Standard monolithic components with a second source are used as much as possible, with emphasis on CMOS to increase reliability through lower operating temperature and power supply requirements. However for some parts, mostly the SHA but also the open collector bus drivers, the range of available products satisfying system requirements is extremely limited, not to mention the CCD with a specially coated window.

There is one important block which is not shown in the general diagram: the logic that ensures that the microprocessor system is not disturbed while the LCA reads its configuration from the on-board PROM or from the down-load cable. This logic is not perfectly bulletproof but most cases of LCA programming failures do not prevent operation of the microprocessor and other interface boards installed on the bus: the interrupt priority daisy chain remains functional, although attempts to access the board terminate with a bus time-out error.

B.3 Analog part

B.3.1 CCD board

This board is extremely small and simple; its schematic is shown in figure B.2. The driver chips (TSC426) provide the level conversion and high peak current output necessary to drive the ≈ 500 pF input capacitance of the CCD clocks at the required slew rate. Damping resistors control transition times, they also avoid undershoots which inject charge into the substrate and may artificially increase dark current [52, page 230]. The exact value of these resistors may have to be adjusted depending on the specifications of the particular CCD model being used.

The output of the CCD has an amplitude of the order of 2 V at saturation, with a large device to device dispersion. The output voltage is offset by 6 to 7 V with respect to ground. It must then be amplified to match the input scale of the ADC (5 or 10 V) and to drive the cable to the analog board. The most important characteristic of the amplifier is fast settling to 12-bit precision (0.01 %) for gains between 2 and 10. The choice of components and values in the circuit is the result of trials and errors because the effects produced by the load, mostly formed by the capacitance of about one meter of coaxial cable, are difficult to foresee when the objective is to minimize ringing. There is still probably room for improvements although it is likely not to be worth the trouble: tests in obscurity for hours have shown that this circuit exhibits low noise and excellent stability.

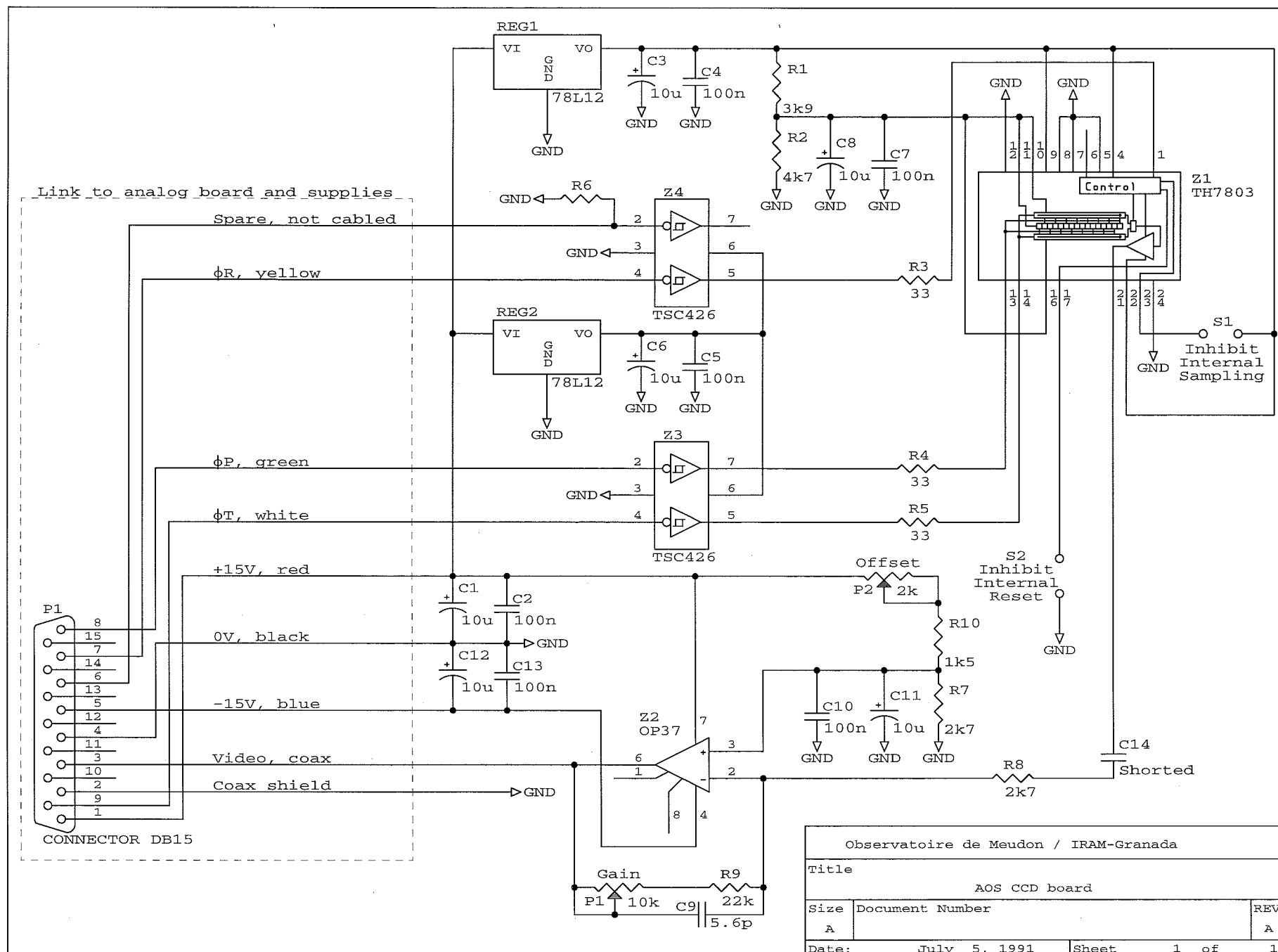
The OP37 amplifier may become unstable and oscillate when the gain is close to the lower limit of the range. If this happens, it can simply be corrected by increasing the gain and using the attenuator at the input of the analog board with the additional benefit that a stronger signal on the video cable improves protection against possible parasitics due to pickup.

B.3.2 Analog board

This board performs the conversion of the amplified CCD output to the digital domain; its schematic is shown in figure B.3.

The technique used for CCD readout is called double correlated sampling and has been

Figure B.2: CCD board schematic.



used for a long time in photon counting applications of CCD. In order to understand how it works, it is necessary to have some knowledge of CCD operation and readout circuitry.

The charge packets in the CCD move towards the readout amplifier under control of the voltage applied to the Φ_T transfer clock. The readout amplifier consists of a back-biased diode which is connected to a reset switch under control of the Φ_R signal and to the gate of the output MOS transistor. On every pixel readout, the diode is reset to a reference level. However, this reset operation is quite noisy and some charges may still remain from the preceding pixel. In order to alleviate this problem, what is actually measured is the difference between the output voltages of the CCD before and after the transition of Φ_T . This technique also has the advantage of rejecting low frequency ($1/f$) noise from the CCD output circuit, which is quite high for MOSFETs. It also rejects flicker noise from the amplifier (which is low anyway) and from the reset reference voltage of the diode. All of this moves the offset and low frequency noise problems after the amplification stage, where they lie at a more manageable level.

The implementation of the double correlated sampling is straightforward, two SHAs respectively hold the reset and signal voltages, and a simple subtractor built around an operational amplifier generates the difference. This difference is then sampled by a last SHA which drives the input of the ADC; the output characteristics, namely impedance and settling time, of this last SHA are critical because of the fast varying load represented by the input of the successive approximation ADC clocked at up to 4 MHz. Another amplifier, connected as a simple voltage follower, buffers the output of the last SHA for monitoring the video signal on an oscilloscope without disturbing operation.

Simple resistor-capacitor networks have been installed at the input of the board and of the last SHA to attenuate broadband noise and reduce feedthrough at high frequencies where the SHAs become somewhat “transparent”. The best compromise between noise and precision is obtained when the filter time constant is close to one tenth of the SHA acquisition time; the 10 to 15 pF parasitic capacitance of each chip input has been taken into account when selecting the values of the capacitors.

The analog board also holds the VFC and some buffers and Schmitt-triggers on the logic input and output signals. There is no signal conditioning circuitry on the VFC input, which is a 20 k Ω resistance to virtual ground, giving a full scale very close to 10 V or 0.5 mA; an input amplifier could easily be added in a next version of the printed circuit if necessary.

Furthermore, the analog board has provision for stand-alone operation: on board logic chips provide the necessary signals to drive the CCD, the SHAs, and the ADC, and jumpers allow to select these signals and to control options of internal logic operation. These chips are not installed in the normal use of the board because the possible feedthrough into the analog electronics of beat frequencies between the internal and external control signals could be extremely harmful. The corresponding circuitry is, for its most part, not shown in figure B.3 and its operation is explained in detail in another document.

B.3.3 Adjustments

Four have been foreseen, two on each board although only the ones on the CCD board are used in practice.

Setting the gain of the video amplifier to optimize the usable dynamic range is by far the most important of all these adjustments. It is easily performed by slightly saturating the CCD and matching the output voltage to the full scale of the ADC. The other trimmer on the CCD board is used to compensate for the DC offset of the CCD video signal; because of

the subtractor in the correlated double sampling circuitry, it must just prevent the output of the amplifier from getting too close to the supply rails and is thus not critical. Attempts to connect the non-inverting input of the amplifier to ground and to AC couple the CCD output to virtual ground gave less satisfactory results, probably because the large variation of the DC offset at the CCD frame frequency makes the common mode rejection ratio of the subtractor in the analog board a critical parameter. A significant decrease of these variations would require an increase in the coupling capacitance to a value only available with electrolytic capacitors.

On the analog board, the first trimmer allows to optimize the common mode rejection of the difference amplifier, i.e., to match the gain of the two paths between the output of the CCD and the difference amplifier. It is however sufficient to use tightly matched precision resistors and to remove the trimmer in the subtractor circuitry. The common mode rejection ratio in this case is not quite as good as with the trimmer adjusted at a given temperature but is significantly less temperature sensitive.

The last adjustment is the offset of the analog part, it acts on the offset pins of the last SHA to prevent the input of the ADC from becoming negative. As we always subtract from the measured spectrum the mean value of the dark current reference pixels of the CCD, the only requirement is that the input voltage of the ADC is never negative even on these pixels. Experimentally, this voltage has always been sufficiently positive to avoid having to compensate for the offset voltages of the circuitry preceding the ADC and the trimmer has been removed on the last versions of the board.

B.4 Digital board

This board has been designed as a standard G96 interface board. Its schematic is shown in figure B.4. The printed circuit is a four layer short (100×160 mm) Eurocard with a solid ground plane giving clean power supplies and signals without power-hungry matching terminations thanks to the small board dimensions.

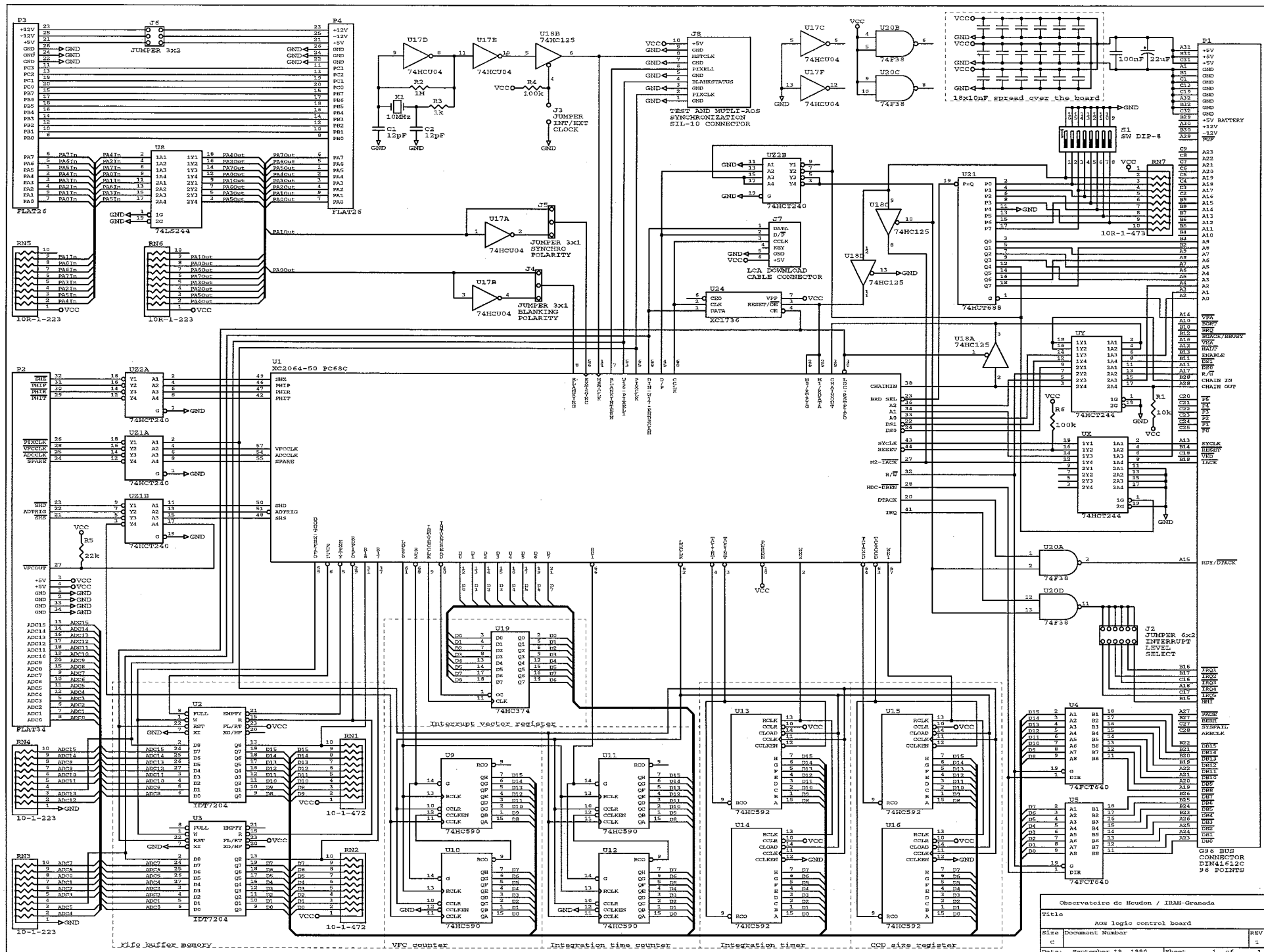
This board can logically be split into two major parts which are separately described in the next sections: one is the sequencer that generates the signals to control the analog and optics boards and the other is the interface with the processor through the G96 bus.

B.5 Sequencer

The sequencer needs to generate all the control signals of the analog part and of the auxiliary external counters. It is the heart and heartbeat of the whole logic and takes up about one half of the LCA. It has been carefully designed for high frequency operation and *glitch-free decoding* of all the signals. Two clocks are used in the design: a high frequency (16 MHz) called *Master clock* that generates the control signals at the pixel level and a Pixel clock derived from the preceding one that generates the control signals at the frame (CCD line) level. Correct operation is guaranteed over the temperature and supply voltage range with the Master clock running up to one third of the LCA specified toggle frequency. This is one of the advantages of the LCA development system, it gives only maximum propagation delays, never typical ones. The resulting designs have enormous timing security margins (a factor 3 is not uncommon) under all but extreme operating environments.

A jumper allows to choose the source of the high frequency clock between an on board crystal oscillator and an external source. The external clock input may be a precise external

Figure B.4: Digital board schematic.



reference frequency or the clock of another spectrometer to make a group of AOS running on the same clock.

B.5.1 Pixel level sequencing

What is described here is the sequence of events occurring between two rising edges of the pixel clock. The processing of one CCD pixel actually spans three such intervals when taking into account sampling of the reference level and storage of the ADC output into the FIFO. A detailed chronogram of all signals, spanning two pixel clock cycles, is shown in figure B.5.

Note that the only signal that may be masked is the FIFO write pulse, all the other signals are generated even after the CCD frame has been completely shifted out. This is done in order to avoid thermal drifts which have been observed on some ADCs when stopping them for several milliseconds before starting another burst of conversions.

It is in theory possible to shorten pixel cycles by overlapping the end of the conversion with the sampling of the reset and signal levels of the next pixel and having the difference ready for sampling by the last SHA. There is however some unavoidable feedthrough and modulating the input of the last SHA affects the ADC by increasing its apparent differential linearity error; experience has shown that it is not worth taking these risks when faster grades of the circuits can provide the required throughput.

The VFC output, which does not appear in figure B.5, typically toggles 200 to 300 ns after the rising edge of the VFC clock.

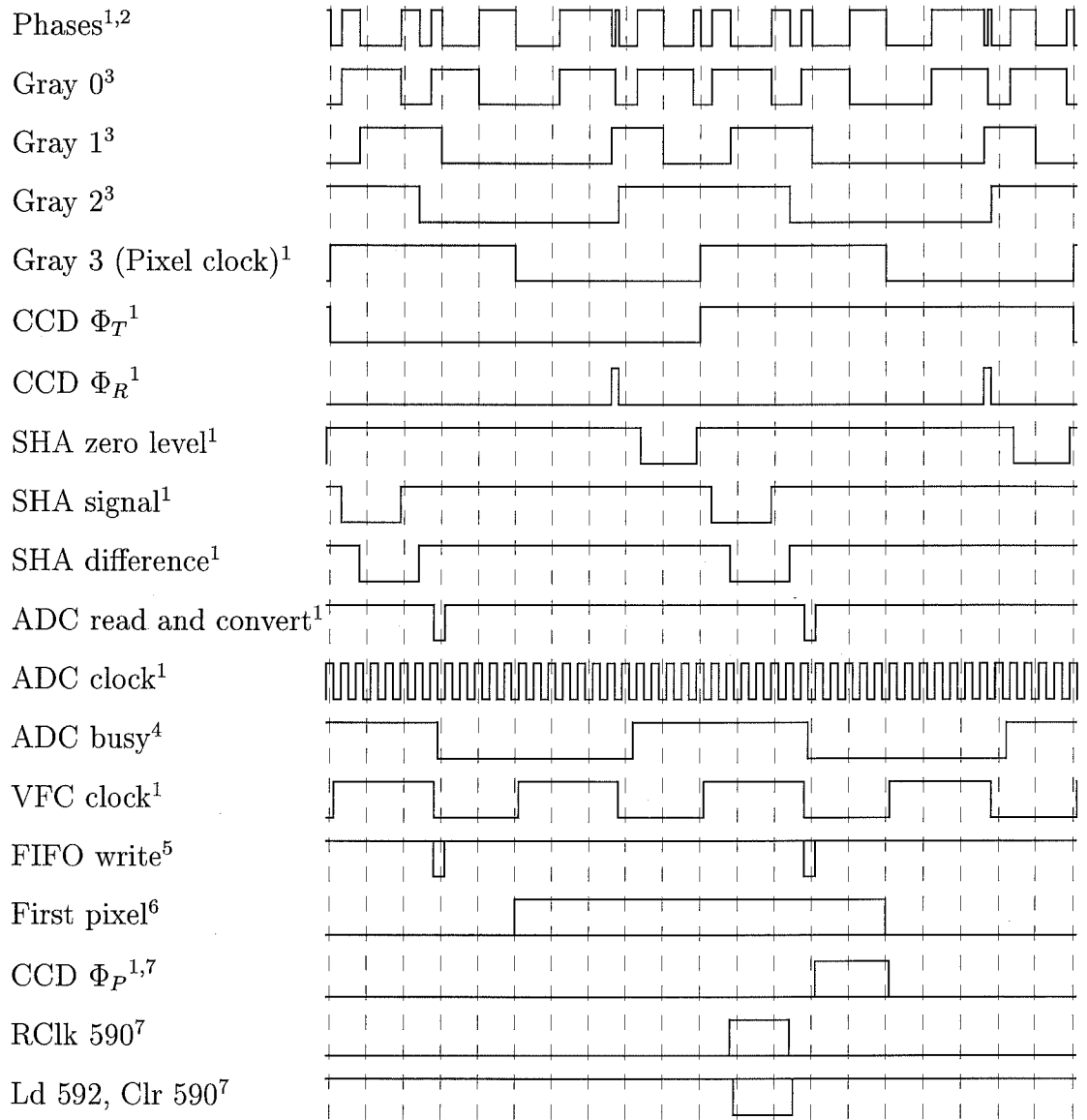
B.5.2 Frame level sequencing

At the frame frequency, the main function of the logic is to decide when it has to trigger a readout, and whether the data has to be written into the FIFO. The chronograms of all important signals are presented in figure B.6 for an hypothetical CCD having at most three active pixels. Some signals which have already been shown on the preceding figure to take advantage of its higher resolution are not repeated here.

These chronograms illustrate the fact that the time between a readout request and the effective generation of CCD Φ_P can be anywhere between 1 and 3 pixel clock cycles, i.e., a delay of up to 20 μ s.

The external blanking signal, when active, asynchronously resets a flip-flop and its complement is clocked into the same flip-flop on the falling edge of First pixel. The state of the blanking logic is frozen at most 100 ns after a request has been accepted, even bus accesses to modify blanking mode are not immediately effective. This ensures that all signals have stabilized when the decision to set or not the Acquire flip-flop, i.e., whether to write data into the FIFO, is taken but implies that the value of the blanking signal is ignored during the last few microseconds of each integration

These delays can safely be ignored with bulky and slow mechanical devices like the wobbler of the 30-m radiotelescope, but should be taken into account when using much faster switching systems like electronically controlled noise sources. In this last case, either the First pixel or the CCD Φ_P signal could be fed-back to the modulation driver electronics to guarantee correct operation.



The interval between two dashed lines represents ten master clock cycles.

¹ This signal is present in inverted form on the flat cable between the analog and digital boards.

² Phases is the exclusive or of the outputs of the Gray code counter and is available on the spare control line for tests.

³ Inaccessible internal LCA signal.

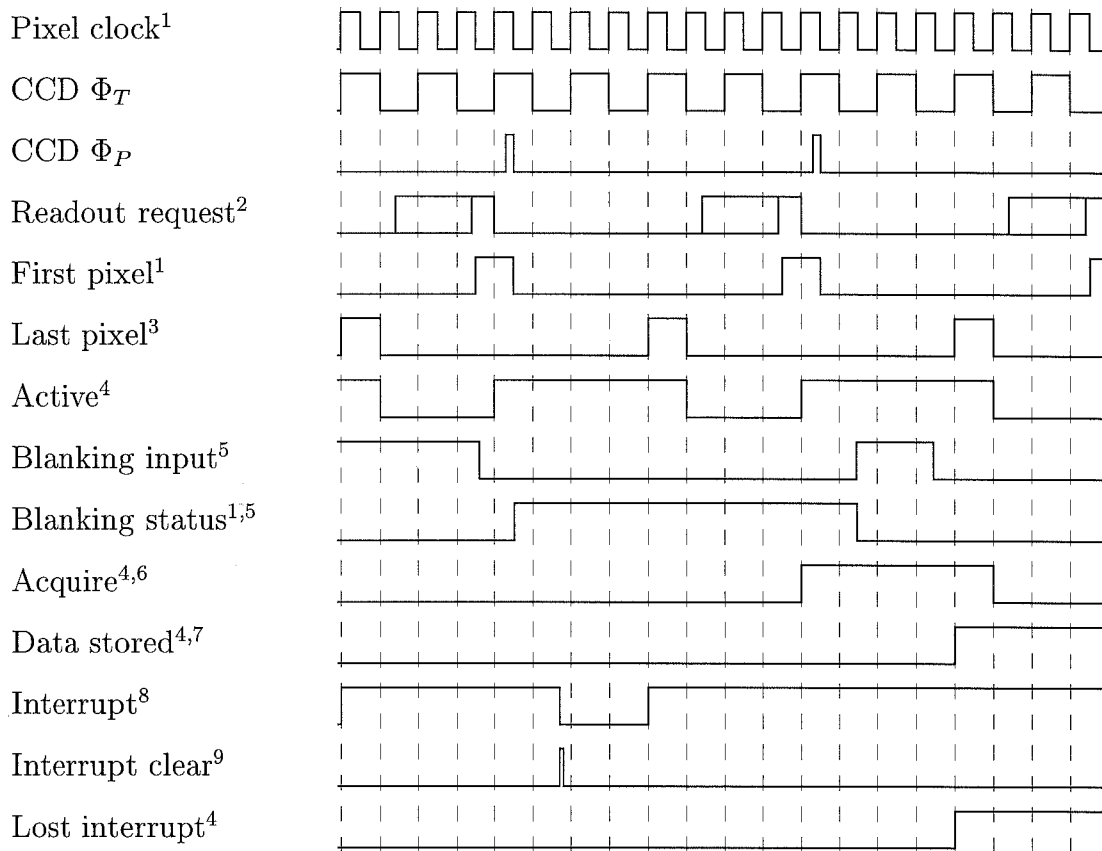
⁴ This signal is only available on the analog board, it is an unused ADC output that is shown only to help checking circuit operation in case of problems.

⁵ FIFO write pulses low only during non blanked CCD readouts.

⁶ Requests for starting a readout are sampled on the high-to-low transition of Pixel clock when Φ_T is low.

⁷ Signal masked by First pixel.

Figure B.5: Pixel level chronogram



The interval between two dashed lines represents one pixel clock cycle, which lasts for exactly one hundred master clock cycles.

¹ This signal is present on the test connector.

² Readout request is an inaccessible internal LCA signal, it is forced to zero as long as Active is high. It may be set by software or, depending on the programmed operation mode, either by an internal timer or by a low-to-high transition on the external readout trigger line.

³ This signal is available during FIFO write pulses on an LCA pin which is otherwise in high impedance with a pull-down resistor.

⁴ The state of this internal LCA signal can be read by software.

⁵ These chronograms assume external blanking is used. In internal blanking control modes, Blanking status merely reflects whether software has enabled writing data into the FIFO. In any case, the value of Blanking status is frozen while Readout request is high. The falling edge of First pixel is used to reinitialize Blanking status to one when appropriate.

⁶ Acquire enables writing into the FIFO. It takes the value of Blanking status on the rising edge of Pixel clock when First pixel is high, it is cleared after CCD readout is complete as indicated by the Last pixel signal.

⁷ Data stored takes the value of Acquire on the rising edge of Last pixel. The interrupt routine can erase it to know whether data has been written into the FIFO.

⁸ The state of this signal can be read by software, it appears on an LCA output and on the bus after inversion only if interrupts are enabled. The first case represents normal operation in which software has time to clear the pending request before the following one is issued, the second one shows the abnormal condition of excessive interrupt response latency causing the Lost interrupt flag to be set.

⁹ Internal LCA signal generated by a software instruction.

Figure B.6: Frame level chronogram

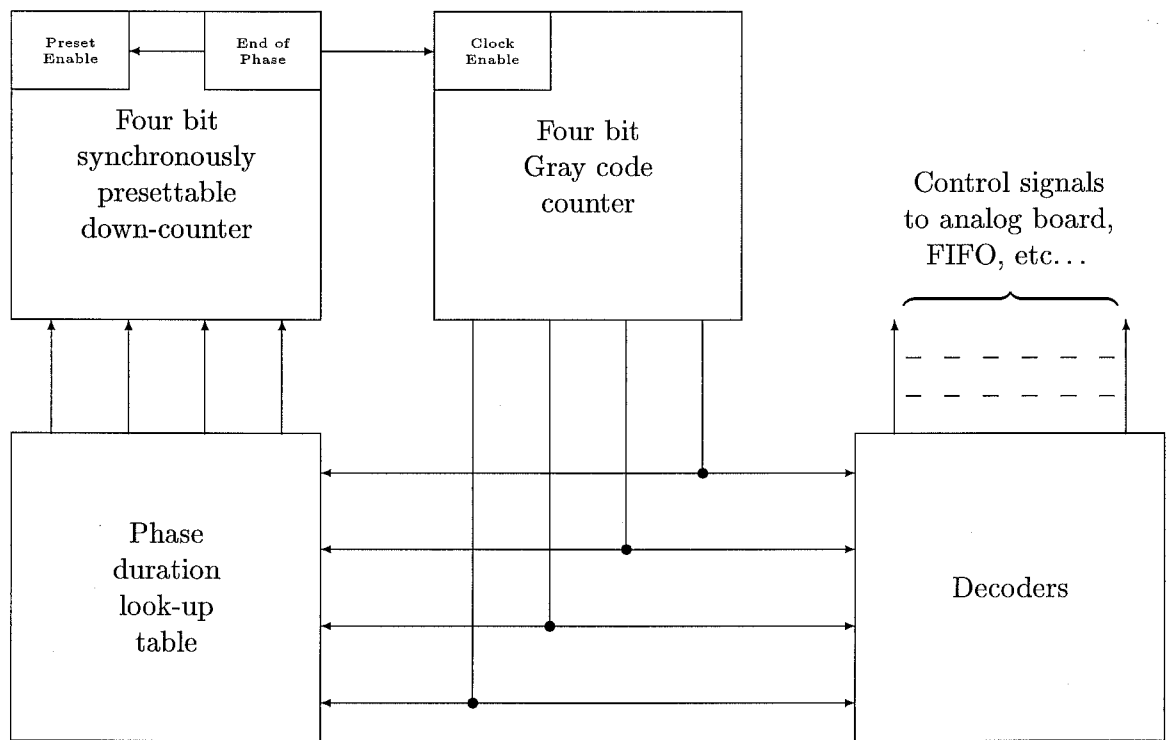


Figure B.7: Sequencer block diagram

B.5.3 Sequencer implementation

The sequencer has been designed to be relatively easy to modify in order to allow upgrades of the analog part without too much effort.

It is desirable to enhance performance and flexibility to have a timing resolution of one master clock cycle. However, as the individual pixel readout cycle lasts a few tens of master clock periods, a brute-force approach with one wide counter would need one six or seven bit decoder per signal, then a synchronization flip-flop to remove possible glitches. This is wasteful of LCA resources, especially because single stage decoders are limited to 4 inputs.

The adopted scheme, shown on figure B.7 takes advantage of the fact that signal changes take place on only a few of the master clock cycles, the only exception being the ADC clock which is generated by dividing the master clock by 4 and synchronized with the ADC convert signal by internal logic. This implies that, for proper operation of the ADC, the frequency ratio between the master and pixel clocks must be a multiple of 4.

The simplest way to explain the inner workings of the decoder is to consider that a pixel readout is divided into 16 phases of duration varying between 1 and 16 master clock cycles. When the first four-bit down counter reaches zero, the second (Gray code) counter is stepped and the first counter is reloaded with the duration of the following phase minus one. The high order bit of the Gray code counter is taken as the Pixel clock signal, which happens to have 50% duty cycle in the current case although this is not a requirement of the logic.

The decoders then only decode the outputs of the Gray code counter. Gray counters have the advantage that only one bit changes between two successive states, and in this case the LCA manufacturer guarantees that no glitches occur at the outputs of the decoders. Another benefit is that signals that last for two phases need only a 3 bit wide decoder, two of which fit inside a single CLB. Most of the control signals are indeed active for two phases

	Gray 0			
	9	10	11	12
	16	15	14	13
Gray 2	1	2	3	4
	8	7	6	5
	Gray 1			
	Gray 3			

Figure B.8: Gray counting sequence on a Karnaugh map.

to take opportunity of this property and further reduce LCA resource usage.

Because of technical constraints, the four bit Gray code counter uses five flip flops, giving 32 possible states. Internal logic ensures that the normal counting sequence is resumed within a few microseconds if one of the 16 incoherent states is accidentally entered.

B.5.4 Modifying the sequencer

Changing the timing may involve modifying either the duration of the phases (the look-up table) or the decoder blocks. Although this presents in theory no difficulty, it is somewhat painful and error-prone because it means juggling with numerous individual CLB configurations and because the Gray counting sequence is not always intuitive. The order of phases is somewhat simpler on the Karnaugh maps which are displayed by the LCA development system, as illustrated by figure B.8 for a four bit decoder, with the convention used in the sequencer that the first phase begins on the raising edge of Pixel clock. Remember however that the entries in the look-up table give the duration of the following phase, and that their values are offset by one because the counter reaches zero before being reloaded. Karnaugh maps are not very useful for 3 bit decoders in their most frequent use of selecting one out of eight states.

In any case, major changes in the timing, like the ones required by a completely new analog board, require careful preparation and some pencil and paper work to first properly draw chronograms and transform them into decoder and look-up table data. However these modifications will hopefully mostly affect the internal logic of CLBs, they will never significantly alter the routing, which, given the limitations and idiosyncracies of the interconnect network, requires much more experience with LCAs to be successfully and efficiently achieved. Once the LCA configuration has been edited, it can easily be downloaded onto a board to check for eventual programming errors with an oscilloscope.

Almost all decoder and all look-up table CLBs have their four logic inputs connected to the four outputs of the Gray-code counter. It may happen that the function performed by one CLB does not depend on all four values; in this case, the design rule checker (DRC) of the LCA development system issues a warning stating that a routed block input is not

used. This message is correct but harmless and the pin should not be removed from its net and unrouted to avoid difficulties when later modifications again require the unused signal.

B.6 Bus interface

This is by far the most annoying part of the system, because it has to blindly follow a standard. Most of the logic is implemented in the lower left edge of the LCA. The large number of signals is at the limit of the LCA capability and achieving routing was quite painful because software was simply unable to perform it automatically, not to say when taking into account timing constraints.

The base address of the board must be a multiple of 10_{16} in the peripheral address space (VPA) of the G96 bus. Its seven high order bits are selected by DIP switch S1.

Typically more than 99 % of the accesses to the board will be FIFO reads. Only FIFO access time may then have a significant performance impact and its timing is optimized by using the fastest possible layout inside the LCA and synchronizing it with the bus clock. The accesses to all the other interface registers have been deliberately slowed down. This permits to design an almost fully synchronous bus control logic using only the master clock, avoiding the possibility of entering incoherent states, for example through metastable transition of a flip-flop. Some asynchronous logic was unfortunately unavoidable, mostly in the area where internal registers control the use of telescope switching status by the sequencer; in this case, signals are first latched in one place a few microseconds before being actually used and are distributed from that unique, synchronized and stable source. The bus interface circuitry is also asynchronously reinitialized at the completion of each access cycle, when it is not only safe but also necessary to do it to ensure that the bus is released within the required delay. This also means that, when following signal paths, the logic may seem to be redundant, especially the generation of the acknowledge and bus buffer enable outputs where the data strobes and board select inputs are used in several places on chained CLBs. This has the advantage that the recovery time of the circuitry is shorter than the minimum interval between bus cycles, imposing no constraints on software.

Whenever the board is accessed or an interrupt acknowledge cycle is started, most of the internal status logic is frozen to avoid undesirable state changes during the access, like a transition of the interrupt request line or of the FIFO empty signal.

The AOS is a real-time device which needs fast interrupt response, for this reason it is connected to the highest priority interrupt line available. Another interrupt level could be used, changing it however not only involves selecting the appropriate bus signal with jumper J2, but also programming its level, which is used during interrupt acknowledge cycles, inside the LCA block indicated on figure B.9. This is extremely easy because no modification of the routing is necessary. If the board is ever used with an autovectorized interrupt line, it is also very simple to disable interrupt vector generation by forcing the output of this block to be permanently zero.

B.7 Power-up and hardware reset

There is an important difference between the two operations, namely that LCA configuration is lost when power is switched off. Hardware reset, although it stops everything, including the sequencer, does not erase LCA contents and operation is resumed as soon as the reset signal becomes inactive. Apart from this, the two terms are strictly synonymous.

One of the switches of S1 allows to choose between configuration from the on board PROM or from the download cable; this last option, while not absolutely necessary, is very convenient and may save lots of memories (which are one-time-programmable) when debugging a new version of the LCA configuration program.

Some pins of the LCA are defined as fixed outputs or inputs from power-up until configuration is complete. The board takes advantage of this feature that constrains somewhat LCA layout because it ties some functions to some specific pins and a few external multiplexers or three-state drivers must be added, while avoiding possible contention. The most important points are detailed here:

1. the D/P (done/program) pin of the LCA, when low, prevents the address decoder from recognizing that the board is being addressed. It also prevents the generation of spurious DTACK (bus data transfer acknowledge) and interrupts, and puts in high impedance the buffers of the A1, bus clock, interrupt acknowledge (another buffer also controlled by D/P multiplexes a configuration mode bit of the LCA on this line), and reset bus lines. This last refinement allows the LCA to ignore reset until it has finished configuration, the power-on reset pulse in most systems lasts long enough to make the configuration delay effectively disappear: software can immediately access the board. More importantly, a reset pulse would abort configuration and restart it without resetting the counters of the PROM, making the configuration impossible to complete and leading to hard to diagnose system failures in some cases.
2. the LCA pin HDC (high during configuration) is used as the active low data bus enable, ensuring no conflict on the bus until the board is operational.
3. internal logic has been designed so that the LCA pin LDC (low during configuration) is also forced low during bus resets. This output is connected to the reset inputs of the FIFOs, emptying them automatically whenever the LCA is reconfigured or reset. This line also enables a buffer that makes the board transparent on the interrupt daisy chain and allows other interface boards to operate even when the LCA is not yet configured; this case often happens during tests when the download cable is being used and would otherwise cause the system to crash or hang.
4. The serial data input pin of the LCA is used as an output during operation, this pin is however used only during the FIFO write pulses and is otherwise three-stated to prevent contention.

B.8 Possible upgrades

The FIFO chips used for the buffer memory present several advantages, mostly due to the fact that address counters are implemented on-chip: a wide and increasing range of storage capacities is available with the same pinout and the simple interface makes it easier to take advantage of the ever decreasing access times. It is thus possible to improve performance or to increase CCD size only by substituting chips. The non random-access characteristic of FIFOs, implying that all data written into must be read out, results in a negligible overhead and is a good architectural match to the serial nature of CCDs.

When the design of the interface board was started, the only existing LCA were the 2000 series; however the specifications of the more powerful and easier to program 3000 series became soon available. Although it was too late to start again the design of the whole logic

with the new generation, the possibility to use the smallest of the new devices, the 3020, has been included since the first version of the printed board because, from the hardware point of view, it is almost identical to the 2064 used in the wire-wrapped prototype board. The only differences involve configuration and are minimal in the master serial and slave modes, which are the only ones used here, not even requiring a switch or a jumper.

It is indeed possible to use any LCA of the 2000 or 3000 series proposed in a 68 pin PLCC package; this now includes the 2064, 2018, 3020 and 3030 models. The increase in complexity afforded by the higher-density chips, up to a factor 2.5 in equivalent gate count claimed by the manufacturer, is probably well beyond what the AOS control logic will ever need.

B.9 LCA layout

An overview of the internal layout of the LCA is given in figure B.9. It shows that almost all resources of the chip are used, including all the IOB and all the pins, because the requirements of the interface were somewhat underestimated when the decision to use a 2064 was taken. Switching to a 3020 would lead to a CLB resource usage of about 70 % and considerably simplify later modifications and improvements, mostly because the two four input functions in each block would ease sequencer design and reduce the number of logic levels in the bus interface logic. It would also be possible to put the interrupt vector register inside the LCA, saving two pins, but using them for other purposes would require a redesign of the printed circuit.

As stated earlier, the only parts which are relatively easy to modify are the decoders, the look-up table and the interrupt level. The locations of the corresponding CLBs are indicated in figure B.9. Any modification of the other parts of the LCA should only be attempted after its operation has been thoroughly understood as there are many intricate details and any change in one place can have unexpected side effects in another area. This is especially true if the routing or logic of the bus interface section is affected.

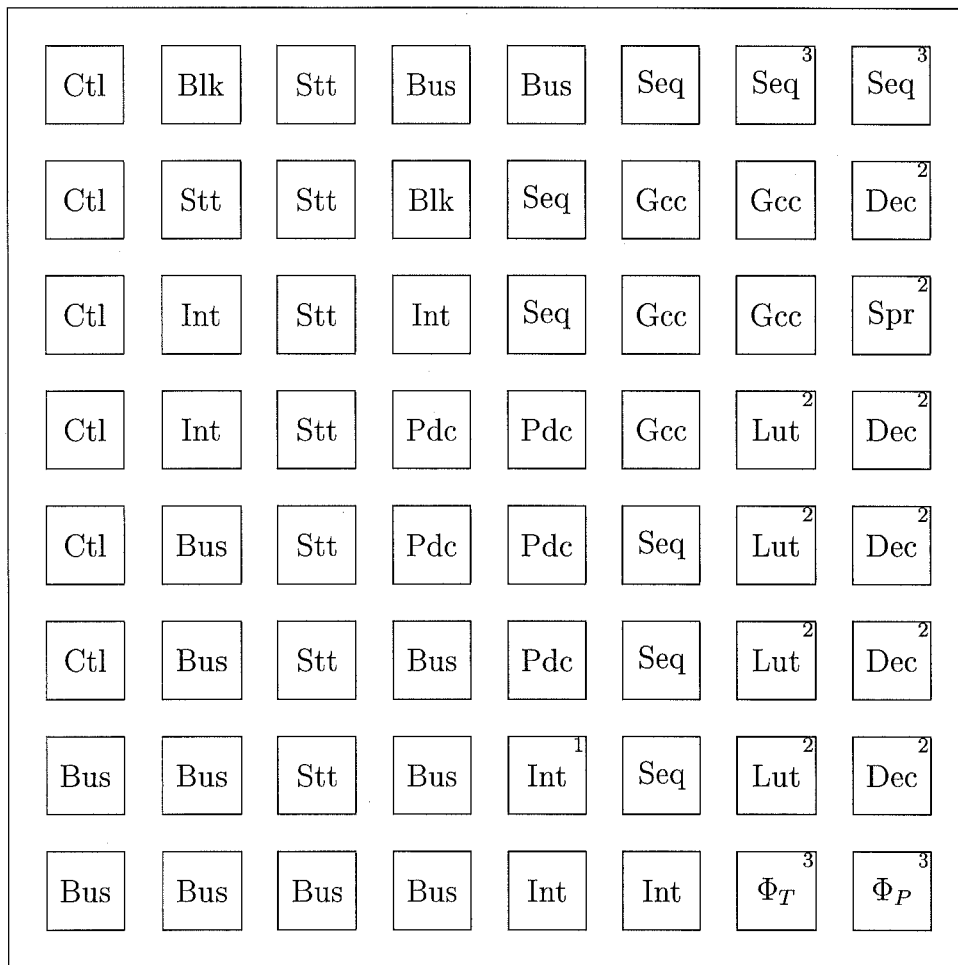
B.10 Programming

This section describes all the capabilities of the interface in term of software instructions, as seen by the programmer that has to write a driver for the AOS.

B.10.1 Address map

It is symbolically shown in figure B.10. The interrupt vector and control and status registers may be accessed as bytes (at an odd address) or words. In the latter case, the upper byte is ignored on writing and undefined on reading. All the other registers must be accessed as words, any attempt to access individual bytes in these registers will terminate with a bus error exception through bus timeout. The FIFO window, being the most often accessed register, is at offset zero so that device drivers can use shorter and faster register indirect instructions.

The first register is the FIFO window, reading it returns the right-justified 12-bit result (i.e., with the four high order bits set to zero) of an analog to digital conversion Readout increments the FIFO internal pointers so that successive reads correspond to successive CCD pixels. Any attempt to read an empty FIFO terminates with a bus error through



Notes:

¹ Interrupt level selection,

² This block may need to be modified when the analog board is updated,

³ Same as note 2, updates may however require routing changes and therefore be much more complex.

Code: Meaning:

Blk blanking logic,

Bus bus interface,

Ctl control register,

Dec control signals decoders (pixel level),

Gcc Gray code counter,

Int interrupt logic,

Lut phase duration look-up table,

Pdc phase duration counter (or presettable down-counter),

Seq frame level sequencer and control signals decoders,

Spr spare decoder (generates the Phases signal in figure B.5),

Stt status registers and multiplexers,

Φ_P , Φ_T corresponding CCD signal.

Figure B.9: LCA layout

Byte offset	High byte	Low byte
00 ₁₆	FIFO(R)	
02 ₁₆	/ / / / / / / /	Interrupt vector(W)
04 ₁₆		Control(W), status ₁ (R)
06 ₁₆		Control(W), status ₂ (R)
08 ₁₆	VFC counter(R)	
0A ₁₆	Time counter(R)	
0C ₁₆	Integration time(W)	
0E ₁₆	Number of pixels per frame(W)	

Figure B.10: Address map of the interface board

bus timeout because no acknowledge is generated, when this happens from supervisor code like the interrupt service routine, it results in a system restart (sometimes referred to as a crash).

The second register holds the interrupt vector byte, which must be loaded by software after power-up and before enabling interrupts. This register is read by the microprocessor during the interrupt acknowledge cycles. It was unfortunately not possible to include logic to check the register contents in the current LCA configuration program.

Third and fourth register control the mode of operation of the system and return various information about its status. Write access to the third and fourth register (byte offsets 5 and 7) have exactly the same effect, while read accesses return different values. These registers consists of single bits or bit fields as shown in figure B.11. They are described in detail in section B.10.2 and B.10.3.

The two following registers (byte offsets 8 and 10) are read-only counters giving respectively the number of VFC output pulses during the last CCD integration and the integration time in units of CCD pixel clock. VFC full scale is exactly equal to the CCD pixel clock frequency, so if the first number is larger than or equal to the second one, it should be interpreted as either a malfunction or saturation.

The last two registers (byte offsets 12 and 14) are write-only. The first one controls the integration time in internal timing mode. The timing unit is, as for the integration time counter register at address 10, the CCD pixel clock. It should be loaded with one minus the desired integration time. However, due to technical constraints imposed by the CCD, the integration time must be an even multiple of the pixel readout time. LCA logic enforces this by waiting for the Φ_T signal to have the correct phase to start the readout. In any case, the exact integration time is always known by reading the register at address 10 after each interrupt.

The second of these registers controls the number of words which are stored into the

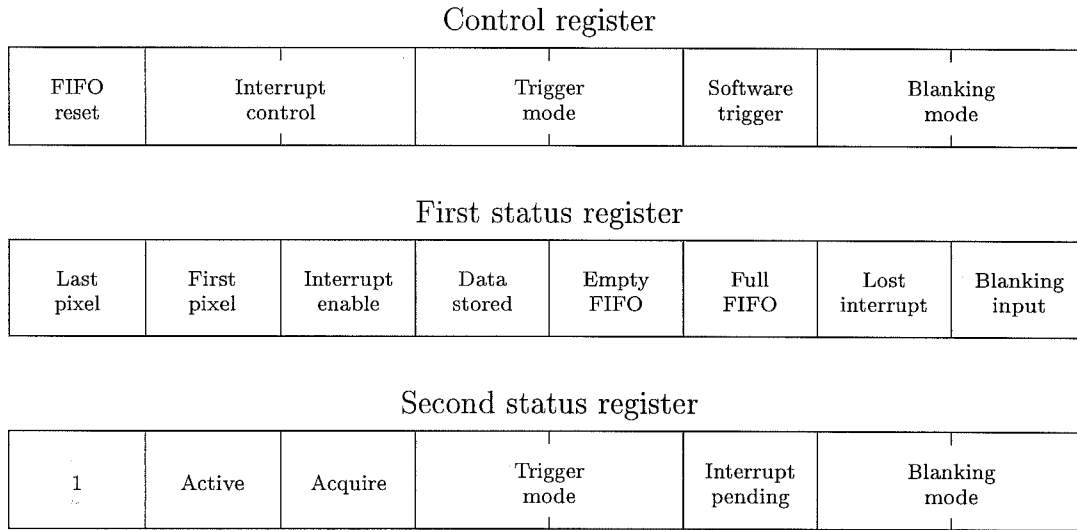


Figure B.11: Control and status registers fields

FIFO for each readout. This number should be at least as large as the number of pixels of the CCD, including the inactive and dark current reference pixels, and smaller than the FIFO capacity. Again, this register should be loaded with one minus the number of words to acquire for each CCD readout.

B.10.2 Operation mode programming

The control register fields are the following, writing a zero in any of the fields has no effect, so that software can simply send orders to the board without having to first read its internal state

FIFO Reset, bit 7: If a one is written into this bit, the FIFO buffer is emptied and the FIFO full flag in status register 1 is cleared. It also clears the Active and Acquire bits, preventing data from being written into the FIFO until the next readout request. This should only be used to recover from catastrophic errors without hardware reset.

Interrupts, bits 5 and 6: Writing 01₂ clears the interrupt pending flag, this should be performed at the beginning of the interrupt service routine. It also clears the Lost Interrupt flag, hence the interrupt service routine should check immediately before that no interrupt has been lost. Writing 11₂ or 10₂ respectively sets or resets the interrupt enable flag. Interrupts are disabled after hardware reset.

Trigger mode, bits 3 and 4: This field selects the way CCD readout is triggered. In order to avoid confusing situations, LCA logic ensures that no readout can start as long as the preceding one has not terminated, which implies that requests occurring too early after the preceding one are ignored. There are three possibilities:

01₂: software mode with integration controlled by the Software trigger bit (see next field). This is the mode after hardware reset.

10₂: internal timer mode, the integration time is given by the integration time register.

11₂: external mode, a rising edge on the external synchronization input of the LCA triggers readout within two pixel clock cycles. The trigger input is connector P3 pin 9 and jumper J5 allows to invert its polarity.

Software trigger, bit 2: Writing a one into this bit triggers CCD readout, even when internal timer or external mode is selected, but only if the Active bit in the second status register is zero.

Blanking mode, bits 0 and 1: There are three possibilities:

10₂: blank always: data writing into the FIFO is disabled. This is the mode after hardware reset.

11₂: write always: data is written into the FIFO regardless of the state of the blanking input. This is used in total power-mode and during tests.

01₂: external blanking: data is written into the FIFO if the blanking input of the LCA has been held low during the CCD integration corresponding to the data being converted, i.e., between the two preceding readout triggers. The blanking signal input is connector P3 pin 7 and jumper J4 allows to invert its polarity. This mode is for example used with the wobbler at the IRAM-30m telescope.

Hardware reset is equivalent to writing CA₁₆ followed by 20₁₆ to clear possible pending interrupts. The only differences lie in the state of bits 4, 6 and 7 of the first status register, which are cleared by hardware reset but can never be changed by software. Performing these two steps exactly as indicated is necessary to guarantee internal consistency after a FIFO overrun and before resuming to normal operation.

B.10.3 Status words

The format of the first status word at byte offset 5 is the following:

Last Pixel, bit 7: is one if the last word read from the FIFO was the last one from a CCD frame.

First Pixel, bit 6: is one if the last word read from the FIFO was the first one from a CCD frame. There are two FIFO chips, each 9 bits wide; the First Pixel and Last Pixel flags are stored as the ninth bit of each FIFO. With the help of this feature, by checking that the correct number of words has been written, software can easily detect problems associated with the FIFOs, which are for example extremely sensitive to parasitic pulses on their read and write inputs. Such a problem, if rare, could otherwise be extremely difficult to diagnose and lead to summing together CCD frames shifted by one pixel. The fact that the two checking bits are not stored on the same chip makes the probability of not detecting a loss of synchronization between the two chips negligible, and greatly increase the confidence in the obtained data. This feature also proved to be extremely helpful while debugging.

Interrupt enable, bit 5: is one if interrupts are enabled.

Data stored, bit 4: is one if the preceding CCD frame has been stored into the FIFO, i.e., if it has not been blanked out.

Empty FIFO, bit 3: when set, indicates the FIFO is currently empty.

Full FIFO, bit 2: when set, indicates the FIFO full output flag has been active. This bit is latched and can only be reset by writing a one into the FIFO Reset control bit or by a hardware reset, it indicates a serious system overload or programming error.

Lost interrupt, bit 1: this flag indicates a severe error or system overload. It means that the interrupt pending flag had not been cleared at the time the following interrupt request was made. The logic generating this flag may in some rare cases fail to detect that the problem is actually occurring, it can however never give false alarms.

Blanking, bit 0: indicates the current logical state of the external blanking input. A one implies that the following readout request will not write data into the FIFO if the blanking mode is external.

The second status register holds the following information, bit 7 is not used and is always one:

Active, bit 6: indicates that a frame is being shifted out of the CCD. It means that readout trigger requests, whatever their origin, are currently ignored.

Acquire, bit 5: indicates that a readout is in progress and that data is being written to the FIFO. Because of different internal propagation delays inside the LCA, it may exceptionnally happen to have the apparently incoherent situation of Acquire being set and not Active.

Synchronization mode, bits 4 and 3: this bit field represents the current synchronization mode. It returns the last code that was written into the corresponding field of the Control register.

Interrupt pending, bit 2: indicates that a CCD readout has completed. The interrupt request line is active if the interrupts are enabled.

Blanking mode, bits 1 and 0: this bit field represents the current blanking mode. It returns the last code that was written into the corresponding field of the Control register.

After hardware reset, the first status register returns 08_{16} or 09_{16} , depending of the state of the external blanking input, and the second one $8A_{16}$.

B.10.4 Interrupts

The AOS electronics has been designed to be used with a dedicated microprocessor giving reasonably fast interrupt response time. As shown in figure B.6, at the end of every CCD readout, whether blanked or not, when the CCD size counter indicates that all pixels have been shifted out, an interrupt request is generated; this request should always be acknowledged and the interrupt pending flag cleared.

The possibility to inhibit hardware interrupts has been given mostly to help while debugging drivers, they should otherwise be permanently enabled even when the AOS is not in use; it is anyway necessary to regularly trigger CCD readout to avoid saturation and the associated long recovery times but the interrupt processing overhead when data is not written into the FIFO is very low, taking less than 1 % of the CPU time. The interrupt routine

can also perform simple and useful tasks like monitoring laser diode power and telescope status signals.

B.10.5 Programming hints

Because it has to handle interrupts, the software must at least partly be written in assembly language. As this also implies that it executes in supervisor mode with lower interrupt levels masked, execution can not be traced with the debugger, making it quite difficult to localize and correct bugs.

The software written for tests in the laboratory and for observations at the IRAM-30m telescope is straightforward, the only noticeable point being that, for performance reasons, the inner loop has explicitly been expanded as shown in program excerpt 1. At the beginning of the loop, the iteration count is divided by 16. The quotient is the loop counter and the remainder is used to jump at the proper location inside the body of the loop for the first iteration.

Apart from the inner loop of CCD readout, the software should mostly be written with security in mind, taking advantages of the features of the logic board. For example, the number of words written to the FIFO is slightly larger than the number of pixels of the CCD. This permits to use the First and Last bits of the status registers to check that the correct number of words have been written and to verify that the readout is properly synchronized without any overhead in the innermost loop of the routine. At the very beginning of the interrupt service routine, it is also recommended to check that neither the Lost interrupt nor the FIFO Full flag is set.

When using the internal timer, a first blanked readout should be triggered by software at power-up or after switching from external trigger to internal timer because the timer counter is in an unknown state at this time and it could take several hundred milliseconds before the timer expires.

B.11 Synchronization between AOS

In some applications, it might be necessary to synchronize several AOS, either to cross-check them by ensuring that they acquire exactly the same data, or to synthesize very wide bandwidths with several spectrographs. In this last case, a careful synchronization may help in avoiding kinks and steps in the baseline, especially under poor weather conditions or when the receiver exhibits gain instabilities. The time lost due to blanking can also be significantly reduced in fast phase switching modes if the switching device can be controlled by the AOS electronics.

Some high speed CCD also have multiple outputs to increase throughput and it may be necessary to have several acquisition boards and even microprocessors to process the data at the required rate.

For these reasons, some hooks for synchronizing AOS have been designed into the hardware. This is implemented as a master-slave relationship between logic boards, with one master and a number of slaves only limited by the capacitive load on the cable linking together all AOS through their test and synchronization connectors. Three of the four signals present on this connector, namely First pixel, Pixel clock and Blanking status are outputs from the master and input to the slave, the last one is the master clock, which can be provided by any board or preferably by an external source with fast edges driving one end


```

* This is the Interrupt Service Routine, all labels begin with ISR.
* a3 points to the AOS interface board throughout the routine.
...
...
* At this point, the number of pixels to read-out
* minus one is in d1.w, and d1.w is nonnegative.
* a0 points to the next pixel in the integration buffer.
* Now, the inner loop is expanded by 16 for speed and is entered
* by a jump into the middle of the loop to perform the correct number
* of iterations; this is admittedly not orthodox structured
* programming but efficient and compact.
* The code for each iteration is four bytes long.
ISRActive moveq    #15,d0    Mask of four low order bits.
* The preceding instruction clears the high order word of D0,
* which stays at zero until the end of the loop.
    eor.w    d0,d1    Complement them.
    and.w    d1,d0    Now d0=15-(d1 mod 16)
    lsr.w    #4,d1    D1= D1 div 16, iteration count for DBRA.
    lsl.w    #2,d0    Multiply by 4 to compute jump offset.
    tst.l    J(a2)    If it's the first readout,
    beq.s    ISRMovJump move and don't accumulate.
    jmp      ISRSumLp(pc,d0.w)
ISRSumLp equ      *      Speed optimized loop
    REPT     16
    move.w   (a3),d0    Read FIFO,
    add.l    d0,(a0)+   and accumulate on 32 bits.
    ENDR
    dbra     d1,ISRSumLp
    bra      ISRCCDLp
ISRMovJump jmp      ISRMovLp(pc,d0.w)
ISRMovLp equ      *      Speed optimized loop
    REPT     16
    move.w   (a3),d0    Read FIFO,
    move.l   d0,(a0)+   and store zero-extended value.
    ENDR
    dbra     d1,ISRMovLp
    ...
    end

```

Programme 1: Excerpt of inner loop code.

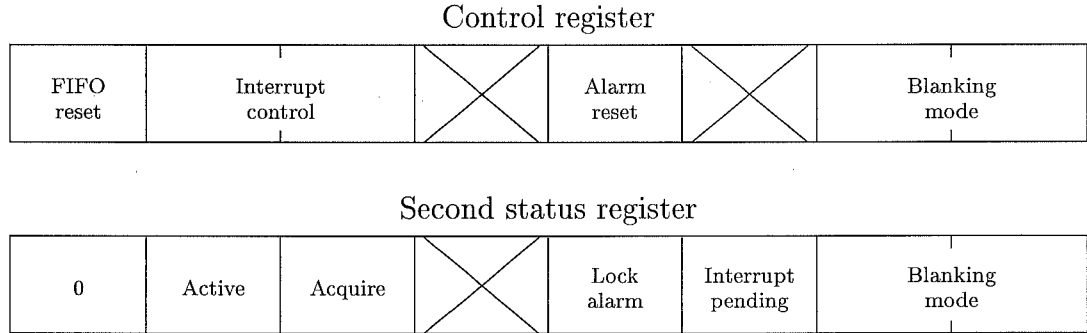


Figure B.12: Modified registers for slave boards

of the cable and suitably terminated at the other end to prevent reflections on this critical high frequency signal.

Apart from running from the same master clock, proper operation requires the following conditions to be fulfilled:

1. all boards have the same pixel clock frequency with the same duty cycle, which means that all analog boards have the same data rate,
2. CCD of different sizes can be used, however the number of pixels of the CCD controlled by a slave board cannot be larger than the one controlled by the master board. More exactly, this restriction applies to the number of words written into the FIFO for each readout.
3. the links between boards for some signals are directly from one LCA to another. Hence, because the inputs of unpowered CMOS devices may present a very low impedance, all boards should run from the same power supply. (Note that the external blanking and synchronization inputs avoid this problem by using a TTL device chosen specifically for this reason)

The internal logic of the slave LCA has been modified so that they get in synchronization with the master. Internally, this is implemented with a circuit which bears some resemblance with a phase lock on the Pixel clock signal. When a phase difference is detected, the phase duration counter is forced to wait for one clock cycle and an alarm flag in a status register is set. This operation repeats until internal and external Pixel clock signals are in phase. CCD Φ_P may however still be 180° out of phase, it is forced into the correct state when First pixel is high. It is therefore necessary to wait for the second readout after power up to ensure that slave boards are properly synchronized.

To support slave mode operation, the definition of some bits in the control and status registers had to be modified as illustrated in figure B.12. The bits controlling and returning the trigger mode have disappeared and replaced by new ones related to the synchronization with the master board. Whenever a hardware reset or a loss of synchronization occurs, bit 2 of the second status register is set; this bit can only be cleared by writing a one into bit 2 of the control register. There are two other differences: bit 7 of the second status register always returns zero, allowing software to know whether the board is master or slave, and external blanking mode takes its input from the blanking status signal generated from the master to guarantee that simultaneous CCD readouts will be either written or not into the

FIFO by all spectrometers. The slave devices in external blanking mode indeed follow the mode of the master.

Extensive tests of spectrometers operating in slave mode have not been performed, however operation without loss of synchronization for several days shows that no major problem should occur and is a good cross-check of the reliability of the sequencer. The measured skew between two boards installed on the same G96 backplane and linked by a short (≈ 30 cm) cable is of the order of 2 ns, another indication of generous timing security margins.

This synchronization logic is a good example of the flexibility of the LCA: it took less than one day to design and implement the slave mode of operation after the decision to give it as an option was taken. It would be much more difficult to implement synchronization in software and it obviously could not have such a low skew.

B.12 Power requirements

The CCD board and the analog board, without chips for stand-alone operation but with local +5 V regulator, together draw about 120 mA and 90 mA, respectively from the +15 V and -15 V supplies, which must have low residual ripple. These figures do not significantly depend on the pixel clock frequency.

The digital board only uses a 5 V supply, it needs 20 to 25 mA when the clock is stopped due to the two bipolar chips, to which must be added 3 mA/MHz of master clock frequency and about 10 mA for bus buffers when acquiring 64 spectras per second with a 1728 pixel CCD. This last term was measured with an 8 MHz processor, it actually decreases when a faster CPU is used simply because bus cycles are shorter.

B.13 Known problems

The first known problem may occur when the power supply is switched off and then switched on again after a very short period of time. It is possible to have the LCA attempting to reconfigure itself without the PROM having reset its internal counters; this depends on the exact value of the power failure detection thresholds in both chips. This was not detected on the prototype board, either because the chips happened not to exhibit the problem, or because the power supply had different startup and shutdown characteristics. There was also a discrete addition in the more recent LCA data sheets specifying a new way to initialize the LCA and that to guarantee that internal power-up reset circuitry is triggered, the supply voltage must be cycled down to below 0.1 V while in the older ones the corresponding value, not explicitly guaranteed, seems to be 2.5 to 3 V. It seems impossible to solve this problem without modifying the printed circuit.

If the microprocessor does not have time to read the integration time and VFC counters between the generation of an interruption and the triggering of the following readout, their values will be lost. This is especially troublesome in external trigger mode where the integration time must be precisely known to normalize the spectras. In order to alleviate this problem, it is possible to modify the LCA configuration to generate interrupts on the First pixel signal. In this case, the values of the counters correspond to the next spectrum and there will be a short additional delay before FIFO readout. These changes should not cause any harm and the required software modifications are quite simple.

There is a third possible problem, although it is unlikely to appear: if the trigger mode is switched from external to internal or software mode between the rising edges of External trigger and First pixel, then the request will stay stored in a flip-flop. If the software then switches back to external trigger mode before any readout has been started, the stored flip-flop will immediately trigger a readout which is indeed not synchronized with the external trigger. It is extremely easy to modify this by letting the required readout proceed and this bug will be corrected in a next update of the LCA configuration.

Appendice C

On the reduction of IRAM-30M Io data

C.1 Introduction

Due to real-time software limitations (and to the lack of ephemeris giving the radial velocities of planetary satellites), the spectra of Io are only corrected for the velocity of Jupiter. Because the lines are much narrower ($< 1 \text{ km.s}^{-1}$) than the change of velocity of Io over one observing night, the spectra have to be resampled on a common grid before summation. This short note attempts to evaluate, under realistic assumptions, the effects of the resampling on the most important parameters of the resulting spectra, like the noise and the equivalent filter shape.

C.2 Notations, Definitions and Hypothesis

The following sections are somewhat technical, and rigorous notations are necessary to avoid confusion. Conventions essentially follow [53] for convenience; only the most important ones are detailed here.

Let $T(\nu)$ denote the antenna temperature of the source at the frequency ν . The individual spectra are defined by the set of n values of the antenna temperature T_i for i between 1 and n . The response of channel i is a function of the frequency ν noted $R_i(\nu)$, $R_i(\nu)$ being non-zero over a finite interval and defined such that

$$\int R_i(\nu) d\nu = 1 \quad (\text{C.1})$$

T_i can be expressed as

$$T_i = \int T(\nu) R_i(\nu) d\nu \quad (\text{C.2})$$

this expression and the normalization equation (C.1) are based on the assumption that the calibration process accurately corrects for the differences in gain between channels.

The input and output voltages of the filters are voltages, i.e, real functions of time. $R_i(\nu)$ can be defined as the square modulus of the Fourier transform of the impulse response of the filters [53, pp. 179–180], and is therefore real and even. Hence the domain of integration should include negative frequencies, we will however restrain ourselves to very simple cases in which we can claim that $R_i(\nu)$ is actually $R_i(\nu) + R_i(-\nu)$ and assume that frequencies are always positive, simplifying expressions by eliminating many absolute value bars. The

limits of the integrals are also often not explicitly specified, they are supposed to be “suitably chosen”.

We shall assume that, whatever i and j between 1 and n , R_i and R_j differ by a simple translation of the origin, i.e., there exist two values ν_i and ν_j such that $R_i(\nu + \nu_i) = R_j(\nu + \nu_j) = R(\nu)$ (note that ν is not restricted to be positive here). We can in this case choose the frequencies ν_i as the centroids of the channels, verifying

$$\int \nu R_i(\nu) d\nu = \nu_i \quad (\text{C.3})$$

We shall also assume that the channels are equally spaced, i.e., that we can write $\nu_j - \nu_i = \Delta\nu \times (j - i)$, where $\Delta\nu$ will be called channel spacing. Hence (C.2) becomes

$$T_i = \int T(\nu) R(\nu - \nu_i) d\nu \quad (\text{C.4})$$

In the precise case of electronic bandpass filters, it can also be reasonably assumed that $R(\nu)$ is increasing for $\nu < 0$, decreasing for $\nu > 0$ and is symmetric around 0. However these last hypothesis are absolutely unnecessary in the following derivations, and may not be strictly verified because filters and mixers are likely to have spurious responses, although at such a low level that they can safely be ignored.

The same notations will be used for the resampled spectra, adding a prime to avoid confusion. For example the frequency response of the j^{th} resampled filter will be noted $R'_j(\nu)$.

The astronomical signal can be supposed to be white gaussian noise, and in this case the expected relative power fluctuations of the signal passing through a filter are inversely proportional to the autocorrelation width B_{ac} of the filter [53, pp. 154–155 and 337–338], which can be computed as

$$\frac{1}{B_{ac}} = \int R^2(\nu) d\nu$$

provided R verifies (C.1). The autocorrelation bandwidth of the filters of the 100 kHz backend is ≈ 170 kHz.

The correlation coefficient C_j of filters i and $i + j$ is defined as

$$C_j = \frac{\int R_i(\nu) R_{i+j}(\nu) d\nu}{\sqrt{\int R_i^2(\nu) d\nu \int R_{i+j}^2(\nu) d\nu}} = B_{ac} \int R(\nu) R(\nu - j\Delta\nu) d\nu \quad (\text{C.5})$$

The only special function we will need is the Λ function, defined by

$$\Lambda(x) = \begin{cases} 1 - x, & \text{if } 0 \leq x \leq 1 \\ 1 + x, & \text{if } -1 \leq x < 0 \\ 0, & \text{otherwise} \end{cases}$$

this function clearly verifies (C.1) and is full width at half maximum (FWHM) is 1. Let us evaluate the convolution of an arbitrary real function $f(x)$ with Λ :

$$[\Lambda * f](x) = \int_{-\infty}^{+\infty} \Lambda(t) f(x - t) dt = \int_0^1 (1 - t) [f(x - t) + f(x + t)] dt \quad (\text{C.6})$$

For a normalized function of FWHM equal to δ , replace $\Lambda(x)$ by $\Lambda(x/\delta)/\delta$, noted $\Lambda_\delta(x)$.

C.3 The resampling operation

Before summing, the spectra of Io are individually resampled on a common grid, defined as a set of p frequencies ν'_i , $i = 1 \dots p$ which has the same channel spacing $\Delta\nu$ as the original one. Each channel is replaced by a linear combination of two channels of the original spectrum, so that we can simply express the antenna temperature of the resampled channel i of frequency ν'_i , noted T'_i from the original channels j and $j + 1$, where j is chosen such that $\nu_j \leq \nu'_i < \nu_{j+1}$, as:

$$T'_i = \alpha T_j + (1 - \alpha) T_{j+1}, \quad \text{with } \alpha = \frac{\nu_{j+1} - \nu'_i}{\Delta\nu}$$

the only assumption we will make from now on is that α is uniformly distributed between 0 and 1. This can easily be checked on any set of data with simple CLASS command procedures, and it should be true when summing large numbers of Io spectra at 100 kHz resolution, but is far from obvious for the lower resolution ones.

C.4 Filter shape

From the preceding considerations and definitions, we can obviously write

$$R'_i(\nu, \alpha) = \alpha R_j(\nu) + (1 - \alpha) R_{j+1}(\nu) \quad (\text{C.7})$$

let us transform this expression

$$\begin{aligned} R'_i(\nu, \alpha) &= \alpha R(\nu - \nu_j) + (1 - \alpha) R(\nu - \nu_j - \Delta\nu) \\ &= \alpha R(\nu - \nu'_i + (1 - \alpha)\Delta\nu) + (1 - \alpha) R(\nu - \nu'_i - \alpha\Delta\nu) \end{aligned}$$

and mean it over α , transforming in the first term $\alpha \mapsto 1 - \alpha$

$$R'_i(\nu) = \int_0^1 (1 - \alpha) [R(\nu - \nu'_i - \alpha\Delta\nu) + R(\nu - \nu'_i + \alpha\Delta\nu)] d\alpha \quad (\text{C.8})$$

this is similar to (C.6), and can indeed be rewritten as

$$R'(\nu) = [\Lambda_{\Delta\nu} * R](\nu) \quad (\text{C.9})$$

Hence the resampled filter is simply the convolution of the original one with a triangle of FWHM equal to the channel spacing. $R'(\nu)$ also verifies (C.1) and (C.3).

If the filters had a perfectly rectangular passband, a monochromatic line would, after resampling, span 3 channels with coefficients 1/8, 3/4, and 1/8 when the line frequency coincides with the center frequency of a channel, and 2 channels with equal amplitude when the line falls just between two channels.

C.5 Line width

When the line is only a few channels wide, the convolution has some effect on the line width. In this section, we will consider not only the usual full width at half-maximum W_{FWHM} but also the root-mean-square W_{RMS} width defined as

$$W_{\text{RMS}}^2 = \int (\nu - \nu_c)^2 T(\nu) d\nu \quad (\text{C.10})$$

where ν_c is the line center frequency. This quantity is easier to use because W_{RMS}^2 is a variance and hence is additive under convolution [53, pp. 142–143].

The BASE command in CLASS computes W_{RMS} in km.s^{-1} and then multiplies it by $\sqrt{8 \ln 2}$, which is the ratio of W_{FWHM} to W_{RMS} for a gaussian line. The computed value is however very sensitive to noise and to the settings of the window limits; symmetrizing the line apparently gives better results. Also this estimation is evaluated as $\sum (\nu_i - \nu_c)^2 T_i$, which is not exactly the same as (C.10). All these effects are obvious in the case the line spans only one or two channels, as is the case with the 1 Mhz filterbanks. Considering that W_{RMS} is about 40 kHz for the 100 kHz filters and exactly $\delta/\sqrt{6}$ for Λ_δ , we can estimate that W_{RMS} of the resampled 100 kHz filters is roughly 60 kHz.

After symmetrizing, W_{RMS} for the 2 mm line at 143 GHz appears to be close to 160 kHz, for which the correction is slightly above 10 kHz. This clearly indicates that taking into account the filter shape and resampling effects would be important at 3 mm, and may slightly alter the conclusions we draw from the 2 mm line. On the other hand, the 1.3 mm line should not be noticeably affected by the data reduction process.

Lo lines seem to be strongly saturated, which means that the ratio of W_{FWHM} to W_{RMS} is higher than $\sqrt{8 \ln 2}$ (up to $\sqrt{12}$ for a perfectly rectangular line). In this case, Gaussian fits tend to overestimate W_{FWHM} and even more W_{RMS} . Other line profile fits available in CLASS do not give satisfactory results, except perhaps for the SHELL method when the position of the line is fixed. Such fits should anyway only be considered for a fast, preliminary analysis.

C.6 Autocorrelation bandwidth

It is clear from the definition (C.7) of $R'_i(\nu, \alpha)$ that it verifies the same normalization condition (C.1) as R . So the autocorrelation bandwidth for one resampled spectrum is simply

$$\frac{1}{B'_{ac}(\alpha)} = \int \alpha^2 R_j^2(\nu) + (1 - \alpha)^2 R_{j+1}^2(\nu) + 2\alpha(1 - \alpha) R_j(\nu) R_{j+1}(\nu) d\nu$$

as all original filters have the same autocorrelation width B_{ac} and using (C.5), we get

$$\frac{1}{B'_{ac}(\alpha)} = \frac{1 - 2\alpha + 2\alpha^2 + 2\alpha(1 - \alpha)C_1}{B_{ac}}$$

The mean autocorrelation bandwidth is then the inverse of the mean of the preceding expression over α and the final result is

$$B'_{ac} = \frac{3B_{ac}}{2 + C_1}$$

The correlation coefficient of two gaussian filters which cross at the half-maximum points is 0.25. The correlation coefficient of the backend filters is actually somewhat lower, close to 20 %, so it can be estimated that the resulting autocorrelation bandwidth is $\overline{B'_{ac}} \approx 230$ kHz; this last value should be used when computing the theoretical baseline noise of an averaged spectrum.

A similar result should be obtained for the 1 Mhz filterbanks, at least when considering blocks of up to a few tens of channels; the situation is however different for the AOS because adjacent channels in the original spectra are strongly correlated (≈ 70 %). This means that there is only a slight (≈ 10 %) increase in the autocorrelation bandwidth, and also little loss in resolution because the spectra are fully sampled.

C.7 Integrated line intensity

Very often the integrated intensity (sometimes called area) of the line is important because it is the parameter with the lowest uncertainty, and is the only one we can use in case some other line is detected with the 1 MHz filterbanks or the AOS; however, when the line is a few channels wide, precise area measurements depend on the hypothesis that

$$\sum_i \frac{d}{d\nu} R_i(\nu) = 0 \quad (\text{C.11})$$

i being limited to the values for which $R_i(\nu)$ is not negligibly small.

Such an hypothesis is somewhat optimistic with filterbanks, although it is theoretically true with the acousto-optic spectrometer because the line area is simply proportional to the number of deflected photons. It should also be true with correlators.

However the convolution of the channels with the Λ function has some interesting effect: let us evaluate the left-hand side of (C.11) when $R_i(\nu)$ is replaced by $[\Lambda_{\Delta\nu} * R_i](\nu)$, using the fact that “the derivative of a convolution is the convolution of either of the functions with the derivative of the other” [53, page 119]

$$\begin{aligned} \Delta\nu^2 \frac{d}{d\nu} R'_i(\nu) &= \int_{-\Delta\nu}^0 R_i(\nu - t) dt - \int_0^{\Delta\nu} R_i(\nu - t) dt \\ &= \int_{\nu}^{\nu+\Delta\nu} R_i(t) dt - \int_{\nu-\Delta\nu}^{\nu} R_i(t) dt \end{aligned}$$

the sum over i obviously converges to zero for any R function which is non-zero over a finite interval. Hence, provided the uniformity hypothesis on α is verified, possible errors on line area cancel out. This may be important for the 1 MHz filterbanks spectra.

C.8 Correlation between channels

It is obvious that the convolution results in an increase of the correlation between adjacent channels. Let us first evaluate $C'_j(\alpha)$ according to its definition

$$\begin{aligned} C'_j(\alpha) &= B'_{ac}(\alpha) \int [\alpha R_i(\nu) + (1 - \alpha) R_{i+1}(\nu)] [\alpha R_{i+j}(\nu) + (1 - \alpha) R_{i+j+1}(\nu)] d\nu \\ &= \frac{(1 - 2\alpha + 2\alpha^2) C_j + \alpha(1 - \alpha)(C_{j-1} + C_{j+1})}{1 - 2\alpha + 2\alpha^2 + 2\alpha(1 - \alpha) C_1} \end{aligned}$$

Then let us evaluate the correlation coefficients $C'_j(\alpha_1, \alpha_2)$ of the sum of two resampled spectra with respective coefficients α_1 and α_2 , taking into account the fact that, whatever i and j , the correlation of $R'_i(\nu, \alpha_1)$ and $R'_j(\nu, \alpha_2)$ is zero because the two spectra are independant

$$C'_j(\alpha_1, \alpha_2) = \frac{[C'_j(\alpha_1)/B'_{ac}(\alpha_1) + C'_j(\alpha_2)/B'_{ac}(\alpha_2)]}{[1/B'_{ac}(\alpha_1) + 1/B'_{ac}(\alpha_2)]}$$

which means that the correlation coefficients combine like the variances, i.e., that they mean out with weights $\frac{1}{B'_{ac}(\alpha)}$. Finally, we obtain for the mean value

$$C'_j = \frac{4C_j + C_{j-1} + C_{j+1}}{4 + 2C_1} \quad (\text{C.12})$$

As C_0 is always 1, C_1 close to 20% and all others are negligible, only C'_1 and C'_2 are significant and evaluate respectively to 0.41 and 0.05. These figures show that adjacent channels are somewhat correlated but not strongly, i.e., they can still be considered as almost independant.

The goal of this section is to try to find the optimum way of determining the integrated intensity of a line that would only be detected with the low resolution spectrometers. In order to do this, we shall consider the case of a monochromatic line. If the line coincides with the center frequency of a resampled channel, we can either measure the intensity of the central channel, which is 3/4 of the total line intensity, or the area enclosed by a group of three channels. In the first case the uncertainty is equivalent to the rms noise σ of the baseline multiplied by 4/3; in the second case the noise $\sigma_{(3)}$ on the sum of three adjacent channels can be evaluated as

$$\sigma_{(3)} = \sqrt{3 + 4C_1 + 2C_2} \sigma \approx 2.17\sigma$$

In the case the line falls just between two resampled channels and we measure the area under the two center channels, the resulting uncertainty is

$$\sigma_{(2)} = \sqrt{2 + 2C_1} \sigma \approx 1.68\sigma$$

This would indicate that for narrow lines, it is better to center a channel on the line and to measure the intensity of the channel and correcting it for the dilution in the two adjacent channels. Unfortunately, the line width is not negligible compared to the channel spacing, and may be unknown, making the dilution factor (4/3 in the example) higher and its estimations quite hazardous. So the best compromise is to center the line between two resampled channels, provided the hypothesis of uniformity on α is verified.

C.9 Other sources of error

During the observation, the frequency of the local oscillator is regularly updated to take into account the velocity of Jupiter relative to the observatory. The synthesizers used in setting the local oscillator have a finite resolution, giving steps of the order of 30 m.s^{-1} . The corresponding W_{RMS} is slightly less than 10 m.s^{-1} , i.e., 5 kHz at 2 mm, and can obviously be neglected. The same holds for the line broadening due to the phase noise of the local oscillator.

C.10 Conclusion

The preceding sections show that the data reduction of Io spectra must be carefully performed in order to extract all the information from the spectra and to minimize errors on the line parameters.

It seems that the spectra of January 1990 were reduced in a different way, resampling the sum of all preceding spectra before adding the following one; this resulted in strange effects on the noise and highly correlated adjacent channels. This could perhaps explain the fact that the burst seen during only a few minutes was apparently narrower than the global sum.

All synthetic spectra should be computed at very high resolution and later convolved by the appropriate function before being compared with the real ones. This will probably

increase the uncertainty on the results, but they will be closer to the reality. The lines we observe seem to be strongly saturated, but their sharp edges are probably even sharper, and hence the opacity must be even higher.

If we detect a line with the low resolution spectrometers, the choice of the parameters of the resampled spectrum may be critical in improving the precision on the integrated line intensity and its signal to noise ratio.

Liste des Figures

3.1	Modèle de diode micro-onde avec éléments parasites	35
3.2	Schéma de détecteur coaxial dans son circuit d'application.	35
3.3	Caractéristique d'une diode tunnel de détecteur	42
3.4	Schéma du détecteur	47
3.5	Schéma du convertisseur tension fréquence (1)	51
3.6	Schéma du convertisseur tension fréquence (2)	52
3.7	Réponse en puissance sur source de bruit.	56
3.8	Erreur de linéarité différentielle.	57
3.9	Ondulation de la réponse en fréquence du détecteur.	58
3.10	Variance d'Allan à fort niveau	60
3.11	Variance d'Allan à faible niveau	60
3.12	Facteur d'étalonnage en fonction de la puissance.	61
3.13	Dernier pointage avec l'ancien détecteur	63
3.14	et premier pointage avec le nouveau.	63
3.15	Exemple de procédure « ONOFF ».	64
4.1	Modèle classique de l'interaction acousto-optique.	68
4.2	Modèle corpusculaire de l'interaction acousto-optique.	69
4.3	Architecture générale du spectrographe.	75
4.4	Erreur de linéarité sans et avec brouillage	81
4.5	Comparaison du spectrographe acousto-optique avec les batteries de filtres	83
4.6	Forme du filtre équivalent à un canal du spectrographe	84
4.7	Raie du monoxyde de carbone à 230.538 GHz dans IRC+10216	85
4.8	Agrandissement de la figure précédente montrant les pieds de la réponse en fréquence.	85
4.9	Variance d'allan du spectrographe et des filtres	86
4.10	Bande passante sur le premier récepteur à 1.3 mm	87
4.11	Forme de bande utilisée pour mesurer la dynamique	88
4.12	Inverse de la bande de fluctuation en fonction du niveau	89
4.13	Numéro de l'élément central du spectrographe en fonction du temps.	91
4.14	Pas en fréquence du spectrographe en fonction du temps.	91
4.15	Numéro extrapolé de l'élément du DTC correspondant à une fréquence acoustique nulle.	92
5.1	Spectre de Titan dans la première transition de CO	99
5.2	Spectre de Titan dans la deuxième transition de CO	99
5.3	Fonctions de contribution dans la première transition de CO	100
5.4	Fonctions de contribution dans la deuxième transition de CO	101

A.1	3mm SIS receiver on hot load.	107
A.2	3mm SIS receiver on sky.	107
A.3	1.3mm SIS receiver on hot load.	108
A.4	1.3mm SIS receiver on sky.	108
A.5	3mm Schottky receiver on sky.	109
A.6	Noise source on continuum channel 1.	110
A.7	Noise source on continuum channel 2.	110
A.8	Noise source on continuum channel 1.	111
B.1	General diagram of the acquisition system	116
B.2	CCD board schematic.	118
B.3	Analog board schematic.	120
B.4	Digital board schematic.	122
B.5	Pixel level chronogram	124
B.6	Frame level chronogram	125
B.7	Sequencer block diagram	126
B.8	Gray counting sequence on a Karnaugh map.	127
B.9	LCA layout	131
B.10	Address map of the interface board	132
B.11	Control and status registers fields	133
B.12	Modified registers for slave boards	138

Bibliographie

- [1] G. Paubert, D. Gautier, and R. Courtin. The millimeter spectrum of Titan: Detectability of HCN, HC_3N , CH_3CN and the CO abundance. *Icarus*, 60(3):599–612, 1984. Article I.
- [2] A. Marten, D. Gautier, L. Tanguy, A. Lecacheux, C. Rosolen, and G. Paubert. Abundance of carbon monoxide in the atmosphere of Titan from millimeter heterodyne observations. *Icarus*, 76(3):558–562, 1988. Article II.
- [3] L. Tanguy, B. Bézard, A. Marten, D. Gautier, E. Gérard, G. Paubert, and A. Lecacheux. Stratospheric profile of HCN on Titan from millimeter observations. *Icarus*, 85(1):43–57, 1990. Article III.
- [4] E. Lellouch, M. Belton, I. de Pater, G. Paubert, S. Gulkis, and T. Encrenaz. The structure, stability and global distribution of Io's atmosphere. *Icarus*, 98(2):271–295, 1992. Article IV.
- [5] E. Lellouch, J. J. Goldstein, S. W. Bougher, G. Paubert, and J. Rosenqvist. First absolute wind measurements in the middle atmosphere of Mars. *Astrophys. J.*, 383(1):401–406, 1991. Article V.
- [6] J. Crovisier, D. Bockelée-Morvan, P. Colom, D. Despois, and G. Paubert. Observations of parent molecules in comets P/Brorsen-Metcalf (1989o), Austin (1989c1) and Levy (1990c) at millimetre wavelengths: HCN, H_2S , H_2CO and CH_3OH . *ESA Spec. Publ.*, *ESA SP-315*, pages 143–148, 1990. Article VI.
- [7] G. Paubert, A. Marten, C. Rosolen, D. Gautier, and R. Courtin. First radiodetection of HCN on titan. *Bull. American Astron. Soc.*, 19(1):633, 1987.
- [8] A. Marten, D. Gautier, A. Lecacheux, C. Rosolen, and G. Paubert. CO abundance in the stratosphere of Titan from microwave observations. *Bull. American Astron. Soc.*, 19(3):873, 1987.
- [9] B. Bézard, A. Marten, and G. Paubert. First ground-based detection of cyanoacetylene on titan. *Bull. American Astron. Soc.*, 24(3):953, 1992.
- [10] E. Lellouch, G. Paubert, and T. Encrenaz. Mapping of CO millimeter-wave lines in Mars' atmosphere: the spatial variability of carbon monoxide on mars. *Planet. Space Sci.*, 39:219–224, 1991.
- [11] J. Crovisier, D. Despois, D. Bockelée-Morvan, E. Gérard, G. Paubert, L. E. B. Johansson, L. Ekelund, A. Winnberg, W. Ge, W. M. Irvine, W. M. Kinzel, and F. P. Schloerb. A search for the millimetre lines of HCN in comets Wilson 1987 VII and Machholz 1998 XV. *Astron. Astrophys.*, 234(1):535–538, 1990.

- [12] J. Crovisier, D. Despois, D. Bockelée-Morvan, P. Colom, and G. Paubert. Microwave observations of hydrogen sulfide and searches for other sulfur compounds in comets Austin (1989c1) and Levy (1990c). *Icarus*, 93:246–258, 1991.
- [13] J. Crovisier, D. Bockelée-Morvan, P. Colom, D. Despois, and G. Paubert. Observations of parent molecules in comet Austin (1989c1) at millimeter wavelengths: HCN, H₂S, H₂CO and CH₃OH. *Bull. American Astron. Soc.*, 22(3):1089, 1990.
- [14] D. Bockelée-Morvan, P. Colom, J. Crovisier, D. Despois, and G. Paubert. Microwave detection of hydrogen sulphide and methanol in comet Austin (1989c1). *Nature*, 350(6316):318–320, 1991.
- [15] M. Guélin, J. Cernicharo, G. Paubert, and B. E. Turner. Free CP in IRC+10216. *Astron. Astrophys.*, 230:L9–L11, 1990.
- [16] J. Cernicharo, C. A. Gottlieb, M. Guélin, T. C. Killian, G. Paubert, P. Thaddeus, and J. M. Vrtilek. Astronomical detection of H₂CCC. *Astrophys. J., Lett.*, 368(2):L39–L41, 1991.
- [17] J. Cernicharo, G. Paubert, G. Butin, S. Sanchez, H. Steppe, S. Liechti, and W. Brunswig. Astronomical tests of the new distribution box for the 30-m radiotelescope, July 1990.
- [18] D. Morris, D. John, P. Garcia, G. Butin, G. Paubert, J. Peñalver, F. Urbano, and C. Thum. Holography of the 30 meter telescope in August 1990.
- [19] A. Lecacheux, C. Rosolen, P. Dierich, and G. Paubert. Acousto-optical spectrometers for broadband millimeter radioastronomy at I.R.A.M. *Int. J. Infrared Millimeter Waves*, 14(2):169–184, 1993.
- [20] J. W. M. Baars, B. G. Hooghoudt, P. G. Mezger, and M. J. de Jonge. The IRAM 30-m millimeter radio telescope on Pico Veleta, Spain. *Astron. Astrophys.*, 175:319–326, 1987.
- [21] J. Cernicharo, C. Kahane, J. Gómez-González, and M. Guélin. Tentative detection of the C₅H radical. *Astron. Astrophys.*, 164:L1–L4, 1986.
- [22] M. Guélin and J. Cernicharo. Astronomical detection of HCCN radical. toward a new family of carbon-chain molecules ? *Astron. Astrophys.*, 244:L21–L24, 1991.
- [23] A. Marten, D. Gautier, T. Owen, D. Sanders, R. T. Tilanus, J. Deane, and H. Matthews. First detection of CO and HCN in the atmosphere of Neptune. *Bull. American Astron. Soc.*, 23(3):1164, 1991.
- [24] H. Steppe, S. Liechti, R. Mauersberger, C. Kömpe, W. Brunswig, and M. Ruiz-Moreno. Millimeter continuum measurements of extragalactic radio sources. II. *Astron. Astrophys., Suppl. Ser.*, 96:441–475, 1992.
- [25] D. T. Emerson, U. Klein, and C. G. T. Haslam. A multiple beam technique for overcoming atmospheric limitations to single-dish observations of extended radio sources. *Astron. Astrophys.*, 76:92–105, 1979.
- [26] S. Guilloteau. Spectral line calibration of the 30-m and 15-m antennas, March 1987.

- [27] A. A. Penzias and C. A. Burrus. Millimeter-wavelength radio-astronomy techniques. *Annu. Rev. Astron. Astrophys.*, 11:51–72, 1973.
- [28] M. L. Kutner and B. L. Ulich. Recommendations for calibration of millimeter-wavelength spectral line data. *Astrophys. J.*, 250:341–348, 1981.
- [29] D. W. Allan. Statistics of atomic frequency standards. *Proc. IEEE*, 54(2):221–230, 1966.
- [30] K. Rohlfs. *Tools of Radio Astronomy*. A & A Library. Springer-Verlag, 1986.
- [31] S. A. Maas. *Microwave Mixers*. Artech House, Norwood, Massachusetts, 1986.
- [32] A. van der Ziel. *Noise in Solid State Devices and Circuits*. John Wiley & Sons, 1986.
- [33] *Electronics Designer's Handbook*. McGraw-Hill.
- [34] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, Oxford, sixth (corrected) edition, 1983.
- [35] N. J. Berg and J. N. Lee, editors. *Acousto-Optic Signal Processing*, volume 2 of *Optical Engineering*. Marcel Dekker, New York, 1983.
- [36] J. F. Rosenbaum. *Bulk Acoustic Wave Theory and Devices*. Artech House, Norwood, Massachusetts, 1988.
- [37] G. M. Borsuk. Photodetectors for acousto-optic signal processors. *Proc. IEEE*, 69(1):100–118, 1981.
- [38] B. N. Khare, C. Sagan, W. R. Thompson, E. T. Arakawa, F. Suits, T. A. Callcott, M. W. Williams, S. Shrader, H. Ogino, T. O. Willingham, and B. Nagy. The organic aerosols of Titan. *Adv. Space Res.*, 4(12):59–68, 1984.
- [39] J. I. Lunine, D. J. Stevenson, and Y. L. Yung. Ethane ocean on Titan. *Science*, 222(4629):1229–1230, 1983.
- [40] E. Lellouch, A. Coustenis, D. Gautier, F. raulin, N. Dubouloz, and C. Frère. Titan's atmosphere and hypothesized ocean: A reanalysis of the *Voyager 1* radio-occultation and IRIS 7.7- μ m data. *Icarus*, 79(2):328–349, 1989.
- [41] D. O. Muhleman, A. W. Grossman, B. J. Butler, and M. A. Slade. Radar reflectivity of Titan. *Science*, 248(4958):975–980, 1990.
- [42] B. L. Lutz, C. de Bergh, and T. Owen. Titan: Discovery of carbon monoxide in its atmosphere. *Science*, 220(4604):1374–1375, 1983.
- [43] Y. L. Yung, M. Allen, and J. P. Pinto. Photochemistry of the atmosphere of Titan: Comparison between model and observation. *Astrophys. J., Suppl. Ser.*, 55(3):465–506, 1984.
- [44] Y. L. Yung. An update of nitrile photochemistry on Titan. *Icarus*, 72(2):468, 1987.
- [45] P. de Laverny. Distribution verticale de l'acide cyanhydrique dans l'atmosphère de titan, Rapport de stage de D.E.A., Juin 1990.

- [46] E. Gérard. The OH radio line in comets: a review. In *Cometary radio astronomy proceedings of a NRAO workshop held at the national radio astronomy observatory, Green Bank, West Virginia*, 1986.
- [47] E. Chassefière and M. Cabane. Stratospheric depletion of CO on Titan. *Geophys. Res. Lett.*, 18(3):467–470, 1991.
- [48] R. R. Howell, D. B. Nash, T. R. Geballe, and D. P. Cruikshank. High-resolution infrared spectroscopy of Io and possible surface materials. *Icarus*, 78(1):27–37, 1989.
- [49] E. Lellouch, M. Belton, I. de Pater, S. Gulkis, and T. Encrenaz. Io’s atmosphere from microwave detection of SO₂. *Nature*, 346(6285):639–641, 1990.
- [50] D. Bockelée-Morvan, J. Crovisier, D. Despois, T. Forveille, E. Gérard, J. Schraml, and C. Thum. A search for parent molecules at millimeter wavelengths in comets P/Giacobini-Zinner 1984e and P/Halley 1982i. *ESA Spec. Publ., ESA SP-250*, 1:365–367, 1986.
- [51] Xilinx, Inc., 2069 Hamilton Avenue, San Jose, CA 95125. *The Programmable Gate Array Design Handbook*, first edition, 1987.
- [52] Fairchild Weston Systems, Inc., 810 W. Maude Ave., Sunnyvale, CA 94086. *CCD Solid State Imaging Technology*, 1987.
- [53] R. N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, second(revised) edition, 1986.

Article I

The Millimeter Spectrum of Titan: Detectability of HCN, HC₃N, and CH₃CN and the CO Abundance

G. PAUBERT, D. GAUTIER, AND R. COURTIN

*Laboratoire d'Astronomie Infrarouge, Observatoire de Meudon, Place Jules Janssen,
92195 Meudon Principal Cédex, France*

Received March 26, 1984; revised July 18, 1984

The flux emitted by Titan's disk in millimeter lines of HCN, HC₃N, CH₃CN, and CO is calculated by means of a radiative transfer formulation which takes into account the sphericity of the atmosphere. It is demonstrated that the plane-parallel approximation for radiative transfer is no longer valid, especially in the core of emission lines, when Titan is not spatially resolved. The antenna temperatures which would be measured by large radiotelescopes observing Titan at frequencies of (1–0) and (2–1) transitions of CO, of (1–0), (2–1), and (3–2) transitions of HCN, and of selected transitions of HC₃N and CH₃CN in the range 80–300 GHz are calculated. The observability of these transitions is investigated. It is concluded that there is the possibility of inferring the vertical stratospheric distribution of these species from line shape measurements to be achieved with existing or forthcoming radioastronomical instrumentation. The determination of the CO abundance by D. O. Muhleman, G. L. Berge, and R. T. Clancy (1984, (*Science (Washington, D.C.)*, **223**, 393–396) from measurements at 115.3 GHz in two 200 MHz bands, is reinterpreted by means of this radiative transfer formulation. A CO mixing ratio between 3×10^{-5} and 18×10^{-5} , with a most plausible value of 7.5×10^{-5} , is found. © 1984 Academic Press, Inc.

I. INTRODUCTION

The Voyager mission has revealed that Titan is characterized by a complex atmosphere, predominantly made of molecular nitrogen, methane, and possibly argon (Broadfoot *et al.*, 1981; Tyler *et al.*, 1981; Samuelson *et al.*, 1981) and containing also hydrocarbons and nitriles in the stratosphere (Hanel *et al.*, 1981; Kunde *et al.*, 1981). Detected nitriles (HCN, HC₃N, C₂N₂) are especially interesting since they are key intermediates in the synthesis of amino acids and bases present in nucleic acids. More recently, CO has also been detected from ground-based observations in the near-infrared (Lutz *et al.*, 1983) and in the millimetric range at 115.03 GHz (Muhleman *et al.*, 1984). Only mean atmospheric mixing ratios have been inferred from these measurements because of insufficient spectral resolution. Moreover, dust and aerosols present in the Titan atmosphere introduce important scattering effects in the radiation transfer even in the

far-infrared range. This renders difficult a proper modeling of the continuum. As a result, the abundances of the various constituents are not very precisely determined (within a factor of 3 at best) which is probably insufficient to test various models of photochemistry on Titan's atmosphere (Samuelson *et al.*, 1983; Strobel, 1982; Yung *et al.*, 1984).

We suggest in the present report that millimeter lines of several atmospheric components on Titan could be detected using modern large radiotelescopes and heterodyne receivers. The vertical distribution of these constituents could be derived if we were able to measure the spectral shape of their emission lines. The interpretation of such measurements would be simpler than in the far-infrared range since scattering is negligible at these wavelengths.

As a first approach, we examine the cases of carbon monoxide (CO), hydrogen cyanide (HCN), and cyanoacetylene (HC₃N) which are known to be present in Titan's stratosphere. We also study the

case of acetonitrile (CH_3CN) which is expected, on the basis of simulation experiments in the laboratory, to be the most abundant nitrile in Titan after HCN and HC_3N (Raulin *et al.*, 1982), and which presents strong spectral transitions in the millimeter range. Atmospheric and spectroscopic data required to calculate the millimeter spectrum of Titan are described in Section II. Due to the relatively large extent of the atmosphere compared to the size of the satellite itself, the plane-parallel approximation for the radiative transfer is no longer valid. A radiative transfer analysis taking into account the sphericity of the atmosphere has been elaborated, as described in Section III. In Section IV, we calculate antenna temperatures which could be measured with a large radiotelescope by observing Titan in HCN, HC_3N , CH_3CN , and CO millimeter lines. In Section V, we discuss the interpretation of the measurement of the (1–0) rotation line of CO by Muhleman *et al.* (1984). In particular, we derive a new estimate of the CO abundance in Titan from our radiative transfer analysis. The detectability of HCN, HC_3N , and CH_3CN millimeter lines by means of forthcoming radiotelescopes and heterodyne receivers is discussed in Section VI.

II. ATMOSPHERIC AND SPECTROSCOPIC DATA

A. Atmospheric Data

The temperature profile derived from the Voyager ingress radiooccultation measurements by Lindal *et al.* (1983) has been used as a mean temperature profile of Titan's atmosphere up to the 0.75-mbar pressure level (corresponding to an altitude of 200 km above the surface). Above this level, the temperature was set constant with altitude and equal to 170°K, and the pressure versus altitude variation was calculated under the assumption of hydrostatic equilibrium at all levels, from the relation:

$$P(r) = P_0 \left(\frac{r_0}{r} \right)^2 \exp \left[\frac{GmM}{RT} \left(\frac{1}{r} - \frac{1}{r_0} \right) \right] \quad (1)$$

where r = the distance from the satellite center,

$P(r)$ = pressure at the distance r ,

r_0 = referencial distance (corresponding to 200 km above the surface),

P_0 = pressure at r_0 ,

G = universal gravitational constant,

R = molecular constant of perfect gases,

m = mean molecular weight,

M = Titan's mass.

We take $GM = 8976 \text{ km}^3 \text{ sec}^{-2}$ from Tyler *et al.* (1981). The mean mixing ratios of the various constituents are given in Table I. In fact, HCN, HC_3N , and CH_3CN , which are supposed to be formed in the upper stratosphere from CH_4 and N_2 photolysis, must condense in the lower stratosphere (Sagan and Thompson, 1984). We assumed that these nitriles are uniformly mixed throughout the stratosphere down to the condensation level indicated on Fig. 1. The partial pressure of each of these constituents follows the saturation law at lower levels down to the tropopause, where it falls to negligible values. CO is not expected to condense for Titan conditions of pressure and temperature. Voyager IRIS data suggest that nitrile abundances vary with latitude and are enhanced in the north polar region. V. Kunde (1984, private communication) indicates the value 2×10^{-7} for HCN at the equator and 2×10^{-6} at the North Pole. For HC_3N , he gives 5×10^{-7} at the North Pole and no upper limit at the equator. On the other hand, CO is probably uniformly distributed on the disk of the satellite. For simplicity, we have assumed a uniform distribution of all considered constituents along the disk. Synthetic spectra shown in this paper were calculated from

TABLE I
TITAN ATMOSPHERIC CONSTITUENTS

Constituent	Detected	Spectral range	Reference	Mixing ratio
HCN	Yes	Far IR (Voyager)	Hanel <i>et al.</i> (1981); Kunde (1984, private communication)	2×10^{-7} (equator) 2×10^{-6} (North Pole)
HC ₃ N	Yes	Far IR (Voyager)	Kunde <i>et al.</i> (1981)	10^{-7} – 10^{-8} (North Pole)
CH ₃ CN	No		Kunde (1984, private communication); expected from laboratory measurements; Raulin <i>et al.</i> (1982)	5×10^{-7} (North Pole) 10^{-7} – 10^{-8}
CO	Yes	Near IR Millimeter	Lutz <i>et al.</i> (1983) Muhleman <i>et al.</i> (1984); revised from this work	2 – 18×10^{-5} 2.3 – 10.2×10^{-5} 3 – 18×10^{-5}

stratospheric mixing ratios of 2×10^{-7} , 1×10^{-7} , and 1×10^{-7} for HCN, HC₃N, and CH₃CN, respectively, but tests were also made with lower abundances, as discussed in Section VI. The value chosen for CH₃CN is based on laboratory simulations made by Raulin *et al.* (1982) which suggest that the abundance of this nitrile on Titan should be comparable to that of HCN. It is also com-

patible with the upper limit of the CH₃CN abundance derived by Cerceau (1983) from the comparison of laboratory spectra recorded at Paris VI University with IRIS North Pole spectra. Similar values were adopted by Sagan and Thompson (1984).

B. Spectroscopic Data

Spectroscopic parameters of HCN, HC₃N, CH₃CN, and CO were taken from the JPL *Submillimeter, Millimeter and Spectral Microwave Line Catalogue* (Poynter and Pickett, 1982). The N₂–N₂ pressure-induced absorption responsible for the continuum opacity was adopted from Courtin (1982). The broadening of spectral lines was represented by Voigt line shape factor at each level. Pressure broadening coefficients were chosen equal to 1.5 GHz atm^{-1} at 300°K , according to Pickett (1982), with a temperature dependence as $T^{-0.5}$, except for CO, for which we adopted a $T^{-0.75}$ dependence from Varanasi and Sarangi (1975).

The frequencies of the transitions exhibited by the four constituents in the 1- to 8-mm-wavelength range are listed in Table II and displayed in Fig. 2, together with the telluric transmission calculated for average conditions of humidity and an altitude of 2400 m. CO and HCN have only two and three lines, respectively, in the range considered, but HC₃N and CH₃CH exhibit much richer spectra. The most appropriate

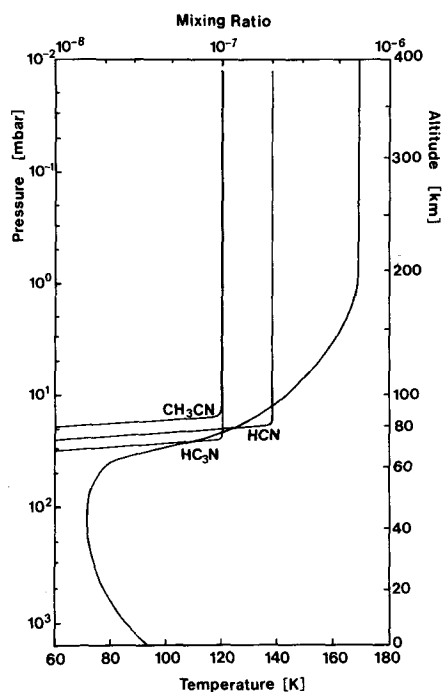


FIG. 1. Temperature profile of Titan and mixing ratios vertical distribution profiles for HCN, HC₃N, and CH₃CN.

TABLE II
FREQUENCY OF MILLIMETER LINES OF CO, HCN, HC₃N, AND CH₃CN

Identification	Frequency (GHz)	Identification	Frequency (GHz)
CO		(22-21)	200.14
(1-0)	115.27	(23-22)	209.23
(2-1)	230.54	(24-23)	218.32
		(25-24)	227.42
HCN		(26-25)	236.51
(1-0)	88.63	(27-26)	245.61
(2-1)	177.25	(28-27)	254.70
(3-2)	265.89	(29-28)	263.79
		(30-29)	272.88
HC ₃ N		(31-30)	281.98
(4-3)	36.39	(32-31)	291.07
(5-4)	45.49		
(6-5)	54.59	CH ₃ CN	
(7-6)	63.69	(2-1)	36.8
(8-7)	72.78	(3-2)	55.2
(9-8)	81.88	(4-3)	73.6
(10-9)	90.98	(5-4)	92.0
(11-10)	100.08	(6-5)	110.4
(12-11)	109.17	(7-6)	128.7
(13-12)	118.27	(8-7)	147.1
(14-13)	127.37	(9-8)	165.5
(15-14)	136.46	(10-9)	183.9
(16-15)	145.56	(11-10)	202.3
(17-16)	154.66	(12-11)	220.7
(18-17)	163.75	(13-12)	239.1
(19-18)	172.85	(14-13)	257.5
(20-19)	181.94	(15-14)	275.9
(21-20)	191.04	(16-15)	294.3

choice of frequencies for detecting Titan emission lines is discussed in Sections IV and VI.

III. RADIATIVE TRANSFER ANALYSIS

In contrast to other planetary atmospheres, the atmosphere of Titan is not negligible in size compared to the Titan body. Due to the large optical path in horizontal viewing conditions, the opacity at the planet's limb can be quite large even at very high altitudes. As a result, the apparent radius of Titan in strong absorption lines is larger than the actual radius of the satellite. The effect is amplified by the fact that the temperature increases with altitude in the stratosphere, which results in a significant limb's brightening. Thus, in the case of

emission lines, an important fraction of the flux emitted by Titan originates from a ring around the planetary disk. An example of this situation is shown in Fig. 3 where the brightness temperature of Titan at several frequencies within the (2-1) line of HCN is calculated as a function of the distance to the center of Titan's disk. In the core of the line, the apparent radius is 15% larger than the actual radius which results in a difference of more than 30% in the total Titan emission. It also appears that the apparent "millimeter" radius of Titan varies significantly and rapidly with frequency within absorption lines, which renders the concept of disk brightness temperature somewhat artificial.

These considerations lead to the conclu-

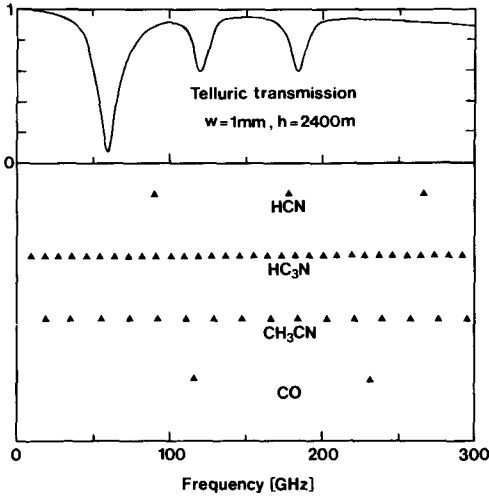


FIG. 2. Top: Telluric transmission up to 300 GHz, at an altitude of 2400 m and for 1 mm of precipitable water vapor (from Zammit *et al.* 1980). Bottom: Positions of transitions for HCN, HC₃N, CH₃CN, and CO up to 300 GHz (see also Table II).

sion that the plane-parallel approximation is not valid in the present case. Therefore the sphericity of the Titan's atmosphere must be incorporated into the radiative transfer formulation, as described below.

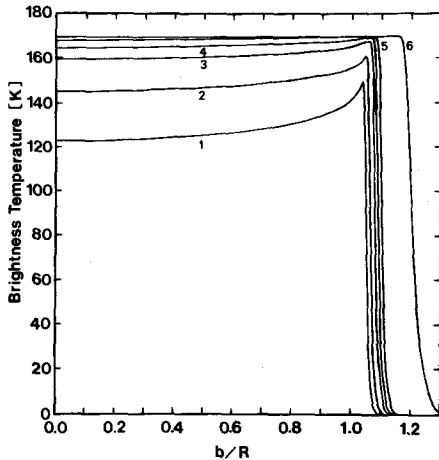


FIG. 3. Variation of the brightness temperature of Titan as a function of distance from the center of the disk. b is the distance from the center and R the radius of Titan. Labels 1, 2, 3, 4, 5, 6 indicate frequencies located at a distance of 81, 31, 11, 6, 3, and 0 MHz respectively, from the central frequency of the (2-1) line of HCN at 177.26125 GHz.

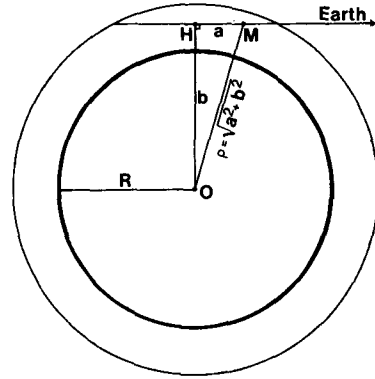


FIG. 4. The geometry of the radiative transfer across Titan's atmosphere. Parameters are defined in the text.

The antenna temperature T_A measured by a radiotelescope is defined by (Kraus, 1966) as

$$T_A = \left(\frac{A}{\lambda^2} \right) \iint T_B(\Theta, \varphi) d\Omega \quad (2)$$

where A = aperture of the telescope,

λ = wavelength of the observation,

$T_B(\Theta, \varphi)$ = brightness temperature of the observed object on the line of sight defined by angular coordinates Θ, φ .

Assuming symmetry of revolution around the Titan-Earth axis (Fig. 4), and introducing the distance D between Earth and Titan, and the impact parameter b defining the distance between the line of sight and the center of Titan, we get

$$T_A = \left(\frac{A}{\lambda^2} \right) \frac{1}{D^2} \int_0^{b_m} T_B(b) 2\pi b db. \quad (3)$$

b_m is chosen in such a way that the optical depth at this altitude is negligible at any frequency. For $b < R$, where R is the radius of Titan, $T_B(b)$ is defined by

$$T_B(b) = T_s \exp[-\tau(0, b)] + \int_{(R^2 - b^2)^{1/2}}^{(b_m^2 - b^2)^{1/2}} T(a, b) \exp[-\tau(a, b)] \gamma(a, b) da \quad (4)^1$$

¹ Neglecting the refraction.

where a represents the position of an atmospheric layer M on the line of sight defined by b (Fig. 4) and $\gamma(a,b)$ is the absorption coefficient within layer M .

$$\tau(a,b) = \int_a^{(b_m^2 - b^2)^{1/2}} \gamma(a',b) da' \quad (5)$$

is the optical depth along the line of sight defined by b . T_s is the surface temperature (assuming a ground emissivity of 1). For $R < b < b_m$

$$T_B(b) = \int_{-(b_m^2 - b^2)^{1/2}}^{(b_m^2 - b^2)^{1/2}} T(a,b) \exp[-\tau(a,b)] \gamma(a,b) da. \quad (6)$$

Numerical calculations have been made using this formulation. The Titan atmosphere was divided in 64 layers extending up to the altitude of 700 km. The Titan radius was taken equal to 2575 km (Lindal *et al.*, 1983).

IV. CALCULATIONS AND RESULTS

Antenna temperatures T_A were calculated for an antenna diameter of 30 m, which corresponds to the size of the French-German-Spanish operated IRAM radiotelescope in Spain, and for a distance Titan-Earth equal to 9.4 AU. Temperature for any other antenna of diameter D can be obtained by multiplying T_A by $(D/30)^2$ where D is expressed in meters.

A. HCN Lines

The three lines of HCN, [(1-0), (2-1), and (3-2)], centered at 88.6, 177.3, and 265.89 GHz, respectively, are optically thick in the atmosphere of Titan, for the abundance values determined from Voyager infrared measurements (see Table I). Therefore the shape of these emission lines reflects the thermal structure of the Titan atmosphere. In the center of the line, the emission originates from warm regions of the upper stratosphere. In the wings of the line, the emission comes from the lower stratosphere where the atmosphere is cooler. Figure 5 illustrates this behavior for the (2-1) line of HCN. The weighting functions (which delineate the atmospheric re-

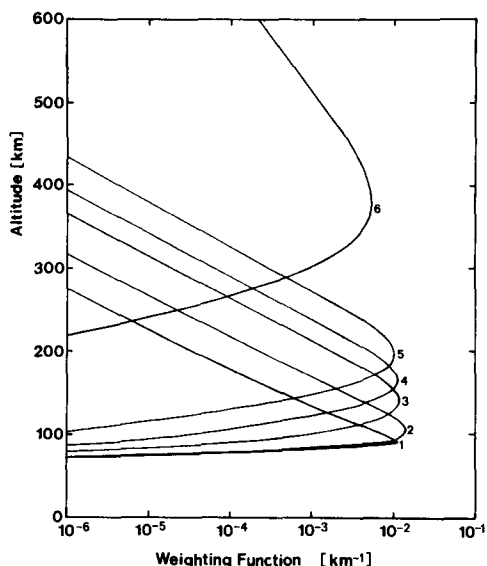


FIG. 5. Weighting functions, as defined by Eq. (7), as a function of altitude at various frequencies within the (2-1) line of HCN at 230.54 GHz. Labels 1, 2, 3, 4, 5, and 6 correspond to frequencies located at 81, 31, 11, 6, 3, and 0 MHz respectively, from the center of the line at 177.26125 GHz.

gions from which the radiation originates) are plotted at a given frequency as a function of altitude.

In the plane-parallel approximation of the radiative transfer, the weighting function $WF_\nu(h)$ is defined by

$$T_B(\nu) = \int_0^\infty T(h) WF_\nu(h) dh \quad (7)$$

where $T_B(\nu)$ is the brightness temperature of the planet for vertical viewing and $T(h)$ is the atmospheric temperature at altitude h . In a first approximation, $T_B(\nu)$ is equal to the temperature of the atmospheric level where the weighting function peaks. Thus, Fig. 5 suggests that information on the thermal structure of Titan could be obtained from 80 to 400 km altitude.

Antenna temperatures which would result from radio observations of Titan by means of a 30-m telescope in the (1-0), (2-1), and (3-2) lines of HCN are shown in Figs. 6-8, respectively. A large dilution factor due to the small angular size of Titan compared to the width of the antenna lobe results in relatively small antenna tempera-

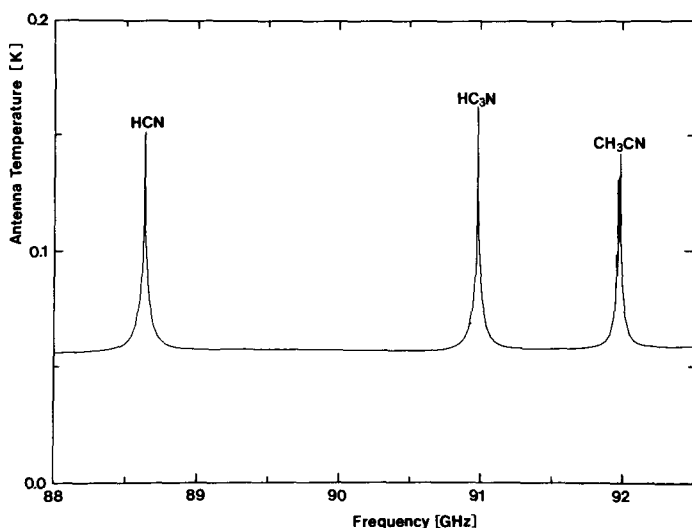


FIG. 6. Antenna temperature spectrum which would be measured by a 30-m-diameter radiotelescope observing Titan in the 88- to 93-GHz frequency range. The (1-0) line of HCN, the (10-9) line of HC₃N, and the (5-4) CH₃CN line are seen in emission.

tures. All three lines clearly appear in emission above the N₂-N₂ continuum. The most intense line arises from the (3-2) transition (Fig. 9) as a consequence of the decreasing width of the antenna lobe with increasing frequency.

B. HC₃N Lines

As seen on Fig. 2, HC₃N has many lines in the millimeter range. We have selected,

as an example, the (10-9) line at 90.98 GHz (Fig. 6), the (24-23) line at 218.32 GHz (Fig. 9), and the (29-28) line at 263.79 GHz (Fig. 8). Note that these lines appear as intense as close-by HCN lines for the mixing ratios used in this work.

C. CH₃CN Lines

CH₃CN also has many lines in the millimeter range but exhibits a much more com-

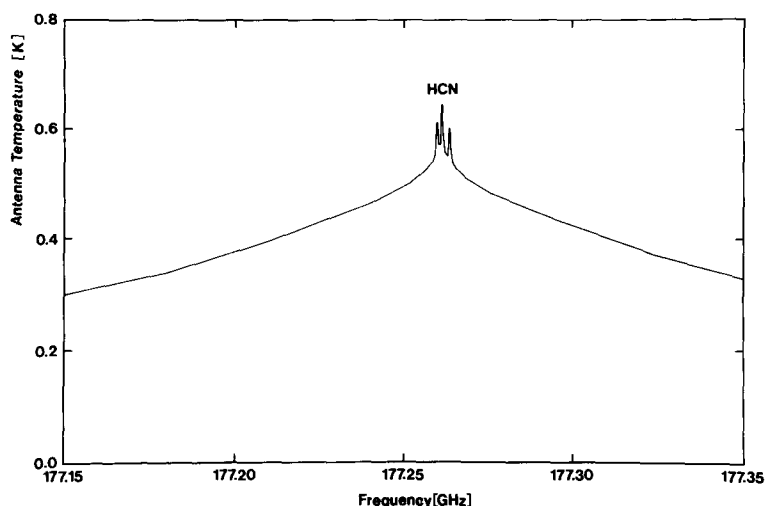


FIG. 7. Same as Fig. 6, around 177 GHz. The (2-1) line of HCN is seen in emission. Note the fine structure of the line.

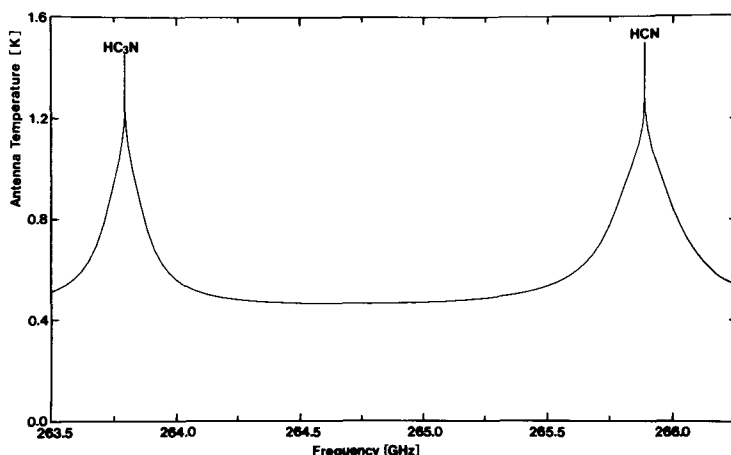


FIG. 8. Same as Fig. 6 in the 263.5- to 266-GHz frequency range. The (3–2) line of HCN and the (29–28) line of HC_3N are seen in emission.

plex structure than HCN and HC_3N . The fine structure is more visible when the order of the transition increases. The (5–4) line centered at 92.0 GHz is shown in Fig. 6 and the (12–11) line centered at 220.7 GHz is shown in Fig. 9.

D. CO Lines

The (1–0) and (2–1) lines of CO centered at 115.27 and 230.54 GHz, respectively, are shown in Figs. 10 and 11. With the mixing ratio of 6×10^{-5} that we have used, accord-

ing to the determination of Lutz *et al.* (1983) from near-infrared ground-based observations, the (1–0) line is twice as intense as the (1–0) line of HCN at 88.6 GHz. It appears to be the easiest line to detect on Titan around 2- to 3-mm wavelength and has indeed been detected in a 200-MHz bandwidth by Muhleman *et al.* (1984), by means of the millimeter-wavelength, two-element interferometer at the Owens Valley Observatory. These authors, using the plane-parallel approximation for the radiative

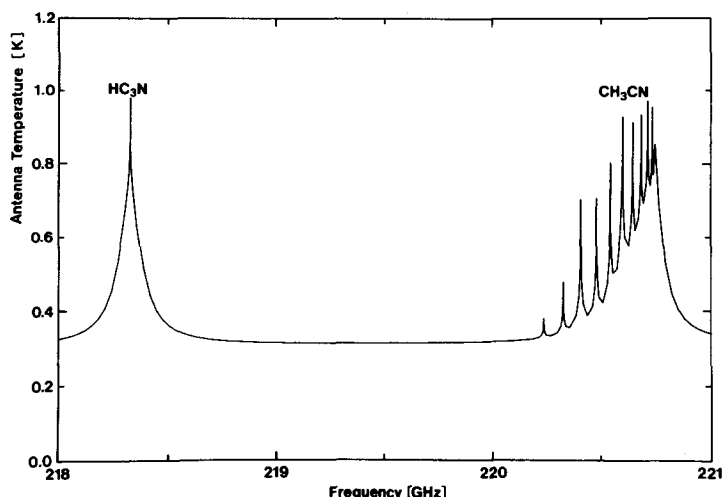


FIG. 9. Same as Fig. 6 in the 218- to 221-GHz range. The (24–23) line of HC_3N and the (12–11) line of CH_3CN are seen in emission. Note the complex fine structure of the CH_3CN line.

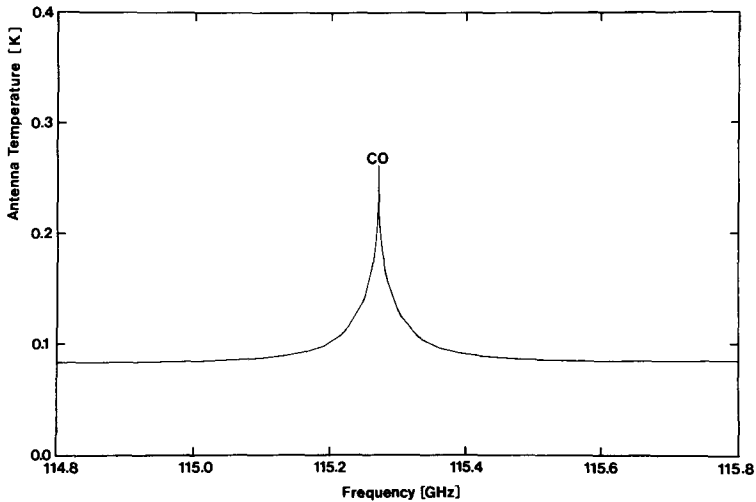


FIG. 10. Same as Fig. 6, around 115 GHz. The (1–0) line of CO is seen in emission.

transfer, have derived a CO mixing ratio between 2.3×10^{-5} and 10.2×10^{-5} . It is interesting to compare these results with the values we can derive from their measurements using the present formulation of the radiative transfer.

V. THE CO ABUNDANCE OF TITAN FROM MICROWAVE OBSERVATIONS

In order to compare our calculations with the results published by Muhleman *et al.* (1984), we have averaged the flux emitted by Titan over frequency in a 200-MHz

bandwidth and we have derived from the total flux the brightness temperature of the disk, assuming, as Muhleman *et al.* did, a Titan radius of 2575 km. The results are shown on Fig. 12 for CO mixing ratios of 2, 6, 10, and 18×10^{-5} . Compared with the results of Muhleman *et al.* for the same mixing ratios (see their Fig. 1), our spectra are warmer by several degrees at all frequencies. We attribute this difference to the fact that our radiative transfer formulation takes into account the atmospheric emission from the limb, which is neglected by

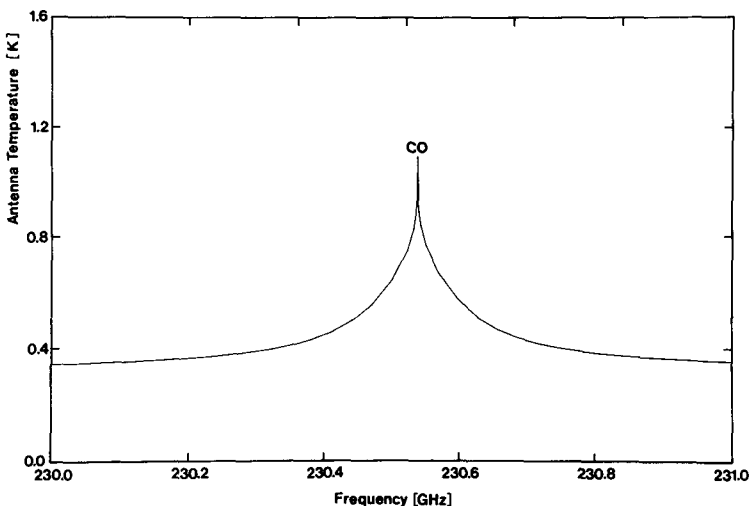


FIG. 11. Same as Fig. 6, around 230.5 GHz. The (2–1) line of CO is seen in emission.

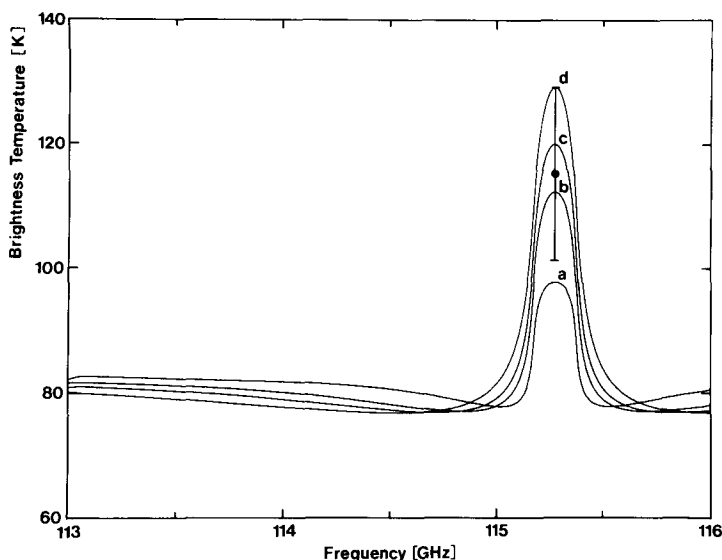


FIG. 12. Synthetic spectrum of the brightness temperature of Titan's disk calculated with a spectral resolution of 200 MHz in the 113- to 116-GHz range. The radius of Titan was taken equal to 2575 km. The (1-0) line of CO appears in emission. Synthetic spectra were calculated for a mixing ratio of CO equal to (a) 2×10^{-7} ; (b) 6×10^{-7} ; (c) 10×10^{-7} ; (d) 18×10^{-7} . Also shown is the measurement of Muhleman *et al.* (1984), revised as described in the text.

Muhleman *et al.* A minor effect comes from the difference in N_2 - N_2 absorption coefficients.² The formulation of Courtin (1982) takes advantage of recent measurements which were not available to Ho *et al.* (1968) when they derived the formulation used by Muhleman *et al.*

The brightness temperature derivation in the CO line ($105 \pm 14^\circ\text{K}$) made by Muhleman *et al.* assumes $T_B = 74.5^\circ\text{K}$ in the continuum at $\nu = 112.6$ GHz. Our calculations indicate that $T_B = 81.5^\circ\text{K}$ at the same frequency. Thus, following the procedure of these authors, we obtain a brightness temperature of $115 \pm 14^\circ\text{K}$ in the CO line, as indicated on Fig. 12. That leads to a CO mixing ratio on Titan between about 3×10^{-5} and 18×10^{-5} , with the most plausible value around 7.5×10^{-5} . The mean value is indeed close to that of Muhleman *et al.*, but the error bars are somewhat larger. Note that Samuelson *et al.* (1983), using a photochemical model and assuming a CO_2 mixing

ratio of 1.5×10^{-9} , predict a constant CO mixing ratio of 1.1×10^{-4} , which is well within the error bars we derived. These uncertainties should be reduced through new measurements with large, forthcoming radiotelescopes, as discussed in the next section.

VI. DETECTABILITY OF CO, HCN, HC_3N , AND CH_3CN MILLIMETER LINES

The comparison of spectra shown on Figs. 6 to 11 clearly shows, as expected, that millimeter emission lines from Titan become more intense as the frequency increases. However, the noise of detectors also increases with frequency so that the most favorable wavelength range for detection of these molecules can be found only by considering both effects.

Typical noise temperatures of receivers T_R in the spectral range of interest as well as the resulting noise level in terms of temperature calculated from

$$\Delta T = \frac{2T_R}{\sqrt{B\tau}} \quad (8)$$

² This effect is, however, mainly responsible for differences in the continuum.

where B is the spectral bandwidth of measurements and τ is the integration time are given in Table III. The factor 2 in (8) takes into account the fact that we are looking for the difference between the emission line and the continuum.

The cores of the emission lines shown on Figs. 6 to 11 can be measured with a bandwidth of about 1 MHz while in the wings of the lines a much larger bandwidth is more appropriate. In Table IV, the values are given of the integration time required to detect the peak of each considered line with a signal-to-noise ratio of 1, 5, and 10 assuming an antenna efficiency of 0.5. A signal-to-noise ratio of 1 implies that the line can be detected by increasing the bandwidth at the expense of the restitution of the line shape. A signal-to-noise ratio of 5 or 10 permits one to measure the line shape, thus providing information on the vertical distribution of the absorber in Titan's atmosphere.

Values of T_R in Table III are in fact very conservative. A number of receivers are now significantly better. For instance, Phillips and Woody (1982) mention SIS receivers routinely used at Owens Valley Radio Observatory for astrophysical observations with a noise temperature of 230°K at 115 GHz. They quote a noise temperature, obtained in the laboratory, of 300°K at 250 GHz. Recently, Blundell *et al.* (1984) have obtained, also in the laboratory, a value of 100°K at 150 GHz. Taking into account present developments in millimeter wave receivers, system noise temperatures of

TABLE III

ν (GHz)	T_R (in °K)	ΔT (°K) for $\tau = 1$ sec	ΔT (°K) for $\tau = 1$ hr
90–115	350	0.7	0.012
150	1000	2	0.035
220–230	2000	4	0.070

Note. ΔT is calculated for $B = 1$ MHz.

TABLE IV

DETECTABILITY OF CO, HCN, HC₃N, AND CH₃CN

Frequency in GHz	ΔT^a (in °K)	$\eta(T_L - T_c)$ (in °K) for $\eta = 0.5^b$	τ for $S/B = 1$	τ for $S/B = 5$	τ for $S/B = 10$
CO					
115.27	0.7	0.08	72 sec	30 min	2 hr
230.54	4	0.32	150 sec	1 hr	4 hr
HCN					
88.63	0.7	0.05	220 sec	90 min	6 hr
177.25	2	0.18	130 sec	55 min	220 min
265.89	4	0.50	64 sec	27 min	105 min
HC ₃ N					
90.98	0.7	0.06	150 sec	1 hr	4 hr
218.92	4	0.35	130 sec	55 min	220 min
263.79	4	0.50	64 sec	27 min	105 min
CH ₃ CN					
92.0	0.7	0.04	5 min	2 hr	8 hr
220.7	4	0.30	3 min	75 min	5 hr

^a ΔT is calculated for $B = 1$ MHz and $\tau = 1$ sec.

^b T_L corresponds to the brightness temperature at the peak of the line, T_c corresponds to the temperature in the continuum, and η is the antenna efficiency.

200°K at 115 GHz, 400°K at 150 GHz, and 1000°K at 250 GHz seem to be realistic now or in the near future, at least for sufficiently dry sites.

Incidentally, it must be noted that the system temperature can become significantly higher when the atmospheric transparency is poor. For instance, the detection of the (2–1) line of HCN at 177.25 GHz is certainly very difficult even under usual atmospheric conditions since this line is in the wing of a strong H₂O telluric line (see Fig. 2). Observations at a low elevation above the horizon also increase the system temperature. A proper evaluation must be made in each case for each site.

In any case, the examination of results given in Table IV permits us to conclude that all lines considered could be easily detected with a 30-m-diameter telescope, provided that the conservative values of detector noise assumed in Table III are achieved.

Results from Table IV also indicate that the shapes of the millimeter lines could be retrieved in a few hours of observation. Except for CO, for which the line at 115.27 GHz is easier to detect than the (2–1) line at

230.54 GHz, high frequency measurements are more favorable for detection. It appears that the detection of these lines could be attempted with smaller telescopes than a 30-m one, especially at high frequencies, since we can compensate to some extent for the growing influence of the dilution factor by an increase in bandwidth. However, another limitation comes from the difficulty in properly evaluating the contribution from Saturn which has an angular diameter of $20''$ and is only distant by about $200''$ from Titan at maximum elongation. In practice, the discrimination of Titan from Saturn requires a telescope of at least 15 m diameter at 90 GHz, or the use of an interferometer.

The observed decrease in nitrile abundances on Titan when going from the North Pole to the equator could make questionable the validity of nominal values of the mean mixing ratios of HC_3N and CH_3CN we have used to calculate disk synthetic spectra. (That is not the case for HCN since we have adopted as a mean value the lower limit 2×10^{-7} corresponding to the abundance determined at the equator). To test the sensitivity of the detected signal to abundance values of HC_3N and CH_3CN , we have redone calculations mixing ratios of HC_3N and CH_3CN we have redone calculations for various mixing ratios values. Results indicate that the peak-to-continuum ratio of HC_3N lines is divided by about 2 when the mixing ratio is divided by 3. The same reduction factor is obtained for CH_3CN when the mixing ratio is divided by 2. In other words, mean stratospheric mixing ratios equal to 3×10^{-8} and 5×10^{-8} for HC_3N and CH_3CN , respectively, would increase by a factor of 4 the integration times required for detection indicated in Table IV. On the other hand, the improvement mentioned above in the receiver's noise temperature will permit to reduce by a factor of 2 the uncertainties ΔT given in Table III, and thus to regain a factor of 4 on integration times. Assuming these improved noise temperatures, it is likely that the shapes of the HCN and CH_3CN lines could

be measured in a realistic time of observation provided that the mixing ratios of these constituents exceed 1×10^{-8} . Weaker lines corresponding to smaller abundances could be detected by increasing the bandwidth. Note that, if HC_3N is, as HCN, 10 times less abundant at the equator than at the North Pole, its mixing ratio at the equator should be around 5×10^{-8} .

Macke (1984, private communication) points out that the actual pressure-broadening coefficients of absorption lines of nitriles are probably significantly larger than the value of 1.5 GHz atm^{-1} we have adopted from Pickett (1982). That is, indeed, strongly suggested by laboratory measurements of CH_3CN in N_2 (Srivastava *et al.*, 1973) and by preliminary results of experiments on HCN in N_2 presently in progress in France, at the University of Lille (Macke, 1984, private communication). However, synthetic spectra calculated with a spectral resolution of 1 MHz are not significantly modified when we increase the broadening coefficients up to 5 GHz atm^{-1} . But the interpretation of measurements made in cores of emission lines with a spectral resolution better than 1 MHz would require the use of precise values for broadening coefficients.

VII. CONCLUSIONS

From synthetic calculations of the millimeter spectrum of Titan developed in this report, we conclude that there is a possibility of measuring the shape of the emission lines of CO, HCN, HC_3N , and possibly of CH_3CN , by using large radiotelescopes and heterodyne receivers. One could then retrieve the vertical distribution of these constituents in the stratosphere of Titan. Such information would lead to a great improvement in the theoretical modeling of the formation process of these molecules.

In the course of this analysis, we have also demonstrated that an accurate calculation of the microwave emission from Titan requires a treatment of the radiative transfer which takes into account the sphericity

of the atmosphere. Reinterpreting the microwave measurements of Muhleman *et al.* (1984) by means of such a formulation, we conclude that the CO mixing ratio in Titan is between 3 and 18×10^{-5} . This uncertainty is too large to provide useful constraints on photochemical models of Titan's atmosphere. On the other hand, the determination of the CO abundance from near-infrared measurements (6×10^{-5}) by Lutz *et al.* (1983) is also uncertain by a factor of 3. The microwave measurements that we propose should permit a significant improvement in the accuracy of the determination of the CO abundance on Titan.

ACKNOWLEDGMENTS

We are grateful to P. Encrenaz, E. Gerard, E. Blum, and L. Wellischew for very useful information on forthcoming radiotelescopes and heterodyne receivers. We are indebted to B. Macke and F. Rohart for their comments and information on pressure broadening coefficients. V. Kunde kindly provided us with recent determinations of nitrile abundances from Titan Voyager spectra. This work was supported by the Institut National d'Astronomie et de Géophysique (ATP Planétologie).

REFERENCES

- BLUNDELL R., J. IBRUGGER, K. H. GUNDLACH, AND E. J. BLUM (1984). Low noise 140–170 GHz heterodyne receiver using quasiparticle tunnel junctions. *Electron. Lett.* **20**, 476–477.
- BROADFOOT, A. L., B. R. SANDEL, E. D. SHEMANSKY, J. B. HOLBERG, G. R. SMITH, D. F. STROBEL, J. C. MCCONNEL, S. KIMAR, D. M. HUNTEN, S. K. ATREYA, T. M. DONAHUE, H. W. MOOS, J. L. BERTAUX, J. E. BLAMONT, R. B. POMPHREY, AND S. LINICK (1981). Extreme ultraviolet observations from Voyager 1 encounter with Saturn. *Science (Washington, D.C.)* **212**, 206–211.
- CERCEAU, F. (1983). *Etude des Spectres Infrarouges de Nitriles Susceptibles d'Être Présents dans l'Atmosphère de Titan*. Thèses, Paris VI University.
- COURTIN, R. (1982). The spectrum of Titan in the far-infrared and microwave regions. *Icarus* **51**, 466–475.
- HANEL, R. A., B. CONRATH, M. FLASAR, V. KUNDE, W. MAGUIRE, J. PEARL, J. PIRRAGLIA, R. SAMUELSON, L. HERATH, M. ALLISON, D. CRUIKSHANK, D. GAUTIER, P. GIERASCH, L. HURN, R. KOPPANY AND C. PONNAMPERUMA (1981). Infrared observations of the Saturnian system from Voyager 1. *Science (Washington, D.C.)* **212**, 192–200.
- HO, W. I., A. KAUFMAN, AND P. THADDEUS (1968). Pressure-induced microwave absorption in N_2 . *J. Chem. Phys.* **49**, 3627–3631.
- KRAUS, J. D. (1966). *Radio Astronomy*. McGraw-Hill, New York.
- KUNDE, V. G., A. C. AIKIN, R. A. HANEL, D. E. JENNINGS, W. C. MAGUIRE, AND R. E. SAMUELSON (1981). C_4H_2 , HC_3N and C_2N_2 in Titan's atmosphere. *Nature* **292**, 686–688.
- LINDAL, G. F., G. E. WOOD, H. B. HOTZ, D. N. SWEETNAM, V. R. ESHLEMAN, AND G. L. TYLER (1983). The atmosphere of Titan: an analysis of the Voyager 1 radio occultation measurements. *Icarus* **53**, 348–363.
- LUTZ, B. L., C. DE BERGH, AND T. OWEN (1983). Titan: Discovery of carbon monoxide in its atmosphere. *Science (Washington, D.C.)* **220**, 1374–1375.
- MUHLEMAN, D. O., G. L. BERGE, AND R. T. CLANCY (1984). Microwave measurements of carbon monoxide on Titan. *Science (Washington, D.C.)* **223**, 393–396.
- PHILLIPS, T. G., AND D. P. WOODY (1982). Millimeter and submillimeter-wave receivers. *Annu. Rev. Astron. Astrophys.* **20**, 285–321.
- PICKETT, H. M. (1982). Microwave spectroscopy in *Vibrational Rotational Spectroscopy for Planetary Atmospheres* (M. J. Mumma, K. Fox, and J. Hornstein, Eds.), NASA Conference Publication 2223, Vol. I, pp. 171–187.
- POYNTER, R. L., AND H. M. PICKETT (1982). *Submillimeter, Millimeter and Microwave Spectral Line Catalogue*. JPL Publication 80-23, rev. 1. Jet Propulsion Laboratory, Pasadena.
- RAULIN, F., D. MOUREY AND G. TOUPANCE (1982). Organic syntheses from CH_4 - N_2 atmospheres: Implication for Titan. *Orig. Life* **12**, 267–279.
- SAGAN, C., AND W. R. THOMPSON (1983). Production and condensation of organic gases in the atmosphere of Titan. *Icarus* **59**, 133–161.
- SAMUELSON, R. E., R. A. HANEL, V. G. KUNDE, AND W. C. MAGUIRE (1981). Mean molecular weight and hydrogen abundance of Titan's atmosphere. *Nature* **292**, 688–693.
- SAMUELSON, R. E., W. C. MAGUIRE, R. A. HANEL, V. G. KUNDE, D. E. JENNINGS, Y. L. YUNG, AND A. C. AIKIN, (1983). CO_2 on Titan. *J. Geophys. Res.* **88**, 8709–8715.
- SRIVASTAVA, G. P., H. O. GAUTAM, AND A. KUMAR (1973). Microwave pressure broadening studies of some molecules. *J. Phys. B* **6**, 743–756.
- STROBEL, D. F. (1982). Chemistry and evolution of Titan's atmosphere. *Planet. Space Sci.* **30**, 839–848.
- TYLER, G. L., V. R. ESHLEMAN, J. D. ANDERSON, G. S. LEVY, G. F. LINDAL, G. E. WOOD, AND T. A. CROFT (1981). Radio science investigations in the

PAUBERT, GAUTIER, AND COURTIN

- Saturn system with Voyager 1: Preliminary results. *Science (Washington, D.C.)* **212**, 201–206.
- VARANASI, P., AND S. SARANGI (1975). Measurements of intensities and nitrogen broadened linewidths in the CO fundamental at low temperatures. *J. Quant. Spectrosc. Radiat. Transfer* **15**, 473–482.
- YUNG, Y. L., M. ALLEN, AND J. P. PINTO (1984). Photochemistry of the atmosphere of Titan: Comparison between model and observations. *Astrophys. J. Suppl. Ser.* **55**, 465–506.
- ZAMMIT, C. C., R. E. HILLS, AND R. W. BARKER (1980). *Atmospheric Emission and Attenuation in the Range 100 to 600 GHz Measured from a Mountain Site*. Internal Report. Mullard Radio Astronomy Observatory, Cavendish Laboratory, University of Cambridge, Cambridge.

Article II

NOTE

Abundance of Carbon Monoxide in the Stratosphere of Titan from Millimeter Heterodyne Observations

A. MARTEN,* D. GAUTIER,* L. TANGUY,* A. LECACHEUX,* C. ROSOLEN,* AND G. PAUBERT†

**Observatoire de Paris-Meudon, 92195 Meudon, France, and †Institute de Radioastronomie Millimétrique, IRAM, Granada, Spain*

Received January 4, 1988; revised April 15, 1988

We searched for the (1-0) rotation line of carbon monoxide on Titan with the IRAM 30-m radiotelescope on Pico-Veleta, Spain. Comparison of the results with previous near-infrared observations of Titan suggests that the stratospheric abundance of carbon monoxide is depleted with respect to the tropospheric value. © 1988 Academic Press, Inc.

In the near-infrared spectrum of Titan, around 1.6 μm , Lutz *et al.* (1983) have clearly identified individual absorption lines of the (3-0) rotation-vibration band of carbon monoxide. The authors derived a CO/N₂ mixing ratio of 6×10^{-5} , with an uncertainty factor of 3 due mainly to the difficulty in properly modeling light scattering by aerosols and possibly by clouds. The observed lines are believed to be formed mainly in the troposphere of the satellite (C. de Bergh, personal communication, 1987). Since CO does not condense on Titan—given the atmospheric temperature conditions—and since it is a very stable molecule, its mixing ratio is expected to be constant with height up to important altitudes.

Subsequently, Muhleman *et al.* (1984) observed Titan in the millimeter range in two 200-MHz bands, one centered on the CO (1-0) line and the other 2.6 GHz away from the line center. The CO emission line is in fact expected to be much narrower than 200 MHz (Paubert *et al.* 1984) so that Muhleman *et al.* could not observe the shape of this feature. The excess flux measured in the CO band compared with the flux in the other continuum band was interpreted by the authors as essentially due to a carbon monoxide emission with an estimate of the most probable CO/N₂ mixing ratio of 6×10^{-5} , assumed constant with height. A reinterpretation of the Muhleman *et al.* data by Paubert *et al.* (1984), who accounted for the emission from the atmospheric limb in the radiative transfer, led to a slightly larger value of the CO mixing ratio, namely $(7.5^{+10.5}_{-4.5}) \times 10^{-5}$. Since the CO millimeter lines are formed mainly in the stratosphere, the agreement between near-infrared and microwave measurements seemed to confirm that CO was uniformly mixed throughout the whole Titan atmosphere.

In this note, we report on heterodyne millimeter observations of Titan made with the IRAM 30-m radiotelescope. Measurements were performed around the frequency of the CO (1-0) line at a spectral resolution of 1 MHz, appropriate for a precise restitution of the expected CO emission lineshape. This information allows, in principle, the inference of the vertical distribution of carbon monoxide in the upper atmosphere of Titan (Paubert *et al.* 1984).

Millimeter observations. Titan was observed on May 5 and 6, 1987, with the IRAM 30-m radiotelescope operating on Pico-Veleta, Spain (Baars *et al.* 1987). Titan's tracking was automatically provided by means of a computer-driven program; the pointing accuracy was better than 3 arcsec rms, as verified by scans through Saturn and Titan itself in broadband continuum measurements. At the dates of the observations, Titan was near an eastern maximum elongation, approximately 160 arcsec away from Saturn and its angular size was 0.75 arcsec. The half-power beam width of the telescope at 115 GHz was about 21 arcsec. Consequently, in the antenna beam, no confusion could exist between Saturn and Titan.

A position switching procedure was used: it consisted of spending half the time on Titan and half the time on an anti-Titan position with respect to Saturn, to further remove any possible contribution from Saturn. For both alternative measurements the scan duration chosen was 30 sec.

We used a 3-mm SIS receiver which operated in the 115-GHz range with a double sideband receiver temperature of about 140°K. The image, local oscillator, and nominal frequencies were at 107.407, 111.430, and 115.271 GHz, respectively. The gain in the nominal band is maintained equal to the gain in the image band.

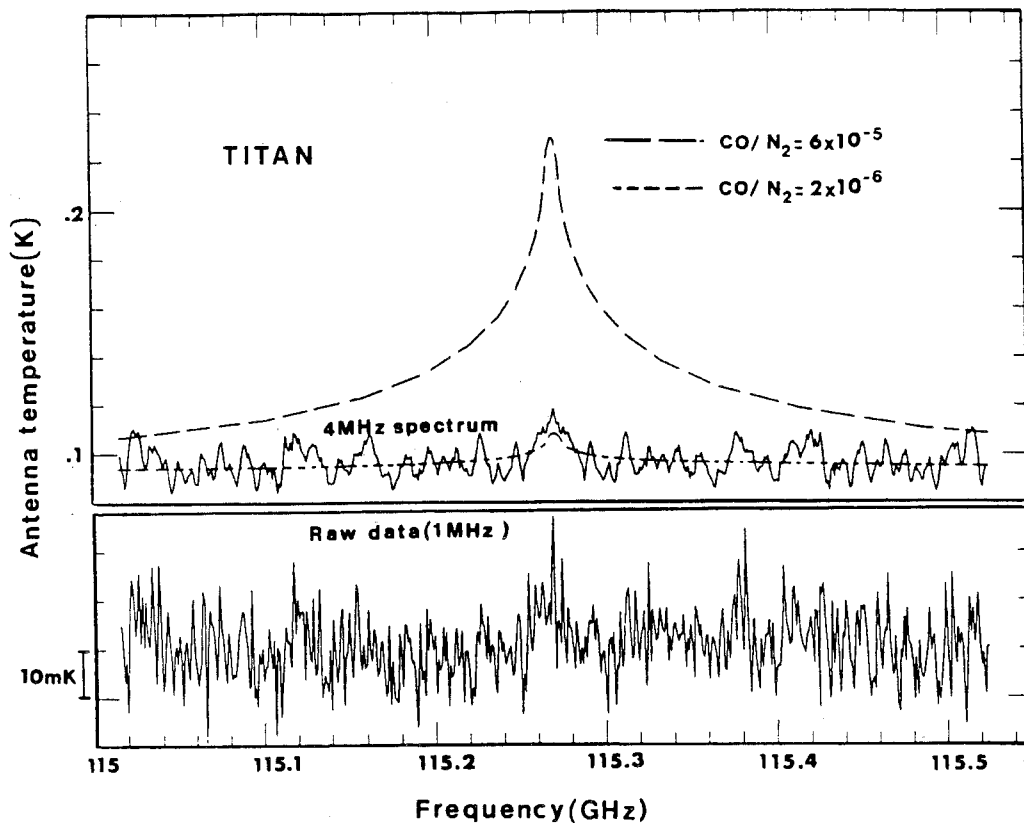


FIG. 1. Observed and synthetic antenna temperature spectra in the (1-0) line of CO centered at 115.27 GHz. At the bottom is shown the raw spectrum recorded with a resolution of 1 MHz, after correction for the antenna efficiency and for telluric opacity. The continuum level is arbitrary but the baseline has not been removed. In the middle, the measured spectrum is shown at a 4-MHz resolution and symmetrized with respect to the central frequency of the CO line, after correction of a 2-MHz Doppler shift due to the orbital motion of Titan around Saturn. The long-dashed line is a synthetic spectrum calculated from a constant-with-height mixing ratio of 6×10^{-5} . The short-dashed line is a spectrum corresponding to a CO/N_2 mixing ratio of 6×10^{-5} up to 60 km, and of 2×10^{-6} at higher altitudes. (See Fig. 2.) The mean value of the received signal has been assigned to the short-dashed line synthetic spectrum, as discussed in the text.

During the observations the total system temperature (referred above the Earth atmosphere, i.e., taking into account the sky emission and corrected for a forward beam efficiency of 0.77) varied between 900 and 1300°K, for an air-mass factor varying between 1.8 and 3.8. The Titan spectrum was analyzed using a back-end system consisting of a 512-channel filter bank, with a channel width of 1 MHz.

Calibrations were regularly made by a chopper wheel method on matched loads at ambient and cold temperatures following the usual procedure adopted for radioastronomical observations at Pico-Veleta (Guilloteau 1987). In addition observations of several reference radio sources [i.e., IRC + 10216, DR21, and SgrB2; cf. Mauersberger *et al.* (1987)] and atmospheric tipping scans were made to accurately verify the internal calibration and to determine the system/receiver temperatures. No spectral baseline ripple effect was seen during calibration sequences. Note that the (1-0) line of HCN was simultaneously observed at 88.6 GHz and that the recorded signal was in excellent agree-

ment with previous observations of Titan we made in September 1986 (Paubert *et al.* 1987).

The raw spectrum of Titan we have thus obtained is presented at the bottom of Fig. 1, in terms of effective antenna temperature above the atmosphere in a 500-MHz frequency range at a resolution of 1 MHz. This spectrum corresponds to a total integration time on Titan of 167 min. The radio continuum of Titan cannot be properly calibrated with the procedure used since it would require a stability of 10^{-4} on both atmospheric transmission and detection system during the on-off sequence. However, the temperature scale, namely the contrast between any emission line and the continuum, is correct with an accuracy better than 10%. Performances of the calibration of lines for the IRAM 30-m telescope have been discussed in detail by Mauersberger *et al.* (1987) and result from a number of radio source measurements made in 1986 and 1987. The noise level was 0.01K rms in antenna temperature.

Discussion. The observed signal is in conflict with

theoretical expectations. The quite small and narrow feature observed at the frequency of the CO line is drastically weaker than the emission line predicted by Paubert *et al.* (1984).

Updated synthetic spectra were calculated following the Paubert *et al.* (1984) radiative transfer technique. The thermal profile used in the calculations was the one inferred by Coustenis *et al.* (1987) from infrared spectra taken within the equatorial region of Titan, recorded by the Voyager 1 IRIS experiment in 1980; this profile has been found to be in good agreement with the Voyager radio-occultation thermal profile of Lindal *et al.* (1983). Spectroscopic parameters of the CO lines were taken from Poynter and Pickett (1982). The CO-N₂ broadening coefficient of the (1-0) line was taken to be equal to 3.33 MHz/Torr at $T = 293^\circ\text{K}$ with a temperature dependence of $T^{-0.86}$, according to Colmont and Monnanteuil (1986). The continuum opacity due to N₂-N₂ pressure-induced absorption was modeled according to Courtin (1982, 1988).

To increase the signal-to-noise ratio of the observed spectrum, the spectral resolution has been downgraded to 4 MHz; moreover, the spectrum was folded, as shown on the middle curve of Fig. 1. As discussed in detail by Gérard (1986), the folding procedure does not introduce any artificial feature.

As mentioned above, the observed continuum level is not well determined. Examination of the shape of the spectrum and comparison with synthetic spectra provide, however, an estimate of at least the upper limit of the CO stratospheric mixing ratio, as discussed below. The synthetic line calculated with a constant-with-height CO/N₂ mixing ratio of 6×10^{-5} , in agreement with near-infrared measurements, exhibits, as shown on the upper curve of Fig. 1, a shape definitely incompatible with our observations, *whatever the continuum level*.

With the present data, we cannot claim firm detection of the (1-0) line of CO but, in any case, the fit of the feature located at the central frequency of the line would require a stratospheric CO abundance of at least an order of magnitude lower than the near-infrared value.

To improve the estimate of the stratospheric CO abundance—in spite of our ignorance of the mean value of the signal—we proceed as follows: synthetic spectra are calculated for various CO mixing ratios; we calculate the mean value M of each spectrum averaged on the whole bandwidth and we constrain the mean of the observed spectrum to be equal to M . The residuals between the so-calibrated observed signal and the synthetic one are subsequently calculated; by trial and error the minimum of the rms of the residuals determines the most plausible CO abundance.

Note that for weak values of the CO mixing ratio, the spectrum at a few tens of megahertz from the line center depends only weakly on the assumed CO distribution since it originates from the troposphere of Titan where the opacity is dominated by the N₂-N₂ pres-

sure-induced absorption. The final result, obtained by trial and error, leads to a stratospheric CO/N₂ mixing ratio of $(2^{+2}_{-1}) \times 10^{-6}$. Expected uncertainties on the thermal profile do not significantly affect this estimate. Even the upper limit (4×10^{-6}) is definitely in conflict with previous CO observations.

We have no explanation for the discrepancy between our measurements and the microwave measurements of Muhleman *et al.* (1984). As previously mentioned, these authors compared the flux density of Titan measured within a 200-MHz band centered at the CO line frequency to the flux measured 2.6 GHz away. The ratio R of the two measurements is 1.48 ± 0.19 . Our synthetic spectrum calculated for CO/N₂ = 6×10^{-5} and integrated over 200 MHz is consistent with this value of R . Other synthetic spectra calculated with much lower CO/N₂ mixing ratios, as low as 2×10^{-6} , lead to values of R smaller than 1.1. However, the two Titan millimeter observations may be statistically compatible at the 2σ level.

On the other hand, our value may not be inconsistent with infrared observations by Lutz *et al.* (1983) since the two observations correspond to different pressure levels sounded in the Titan atmosphere. At millimeter wavelengths, radiation emitted in the CO rotation lines essentially originates in the stratosphere. On the contrary, near-infrared absorption lines, formed mainly in the troposphere (de Bergh *et al.* 1986), provide no information on the lower stratosphere of Titan. It is thus possible to reconcile infrared and microwave observations by assuming that the mixing ratio of the carbon monoxide strongly varies with altitude. By trial and error, it appears that the CO mixing ratio cannot exceed 4×10^{-6} above an altitude of about 60 km (around the 40-mbar pressure level), but below this altitude, it can be anything between 1×10^{-6} and the upper limit 18×10^{-5} of the near-infrared determination. The middle curve of Fig. 1 corresponds to a tropospheric mixing ratio of 6×10^{-5} up to 60 km, and to a mixing ratio of 2×10^{-6} at higher altitudes. This situation, illustrated in Fig. 2, would imply that carbon monoxide is rapidly destroyed, by an unknown process, somewhere around or above the tropopause level.

Such behavior is in conflict with presently available models of the photochemistry in Titan's atmosphere which consider that CO is uniformly mixed throughout the whole atmosphere (Samuelson *et al.* 1983, Yung *et al.* 1984). In the steady-state model published by these authors, the CO abundance depends upon the concentration of CH₃ radicals. A downward revision of the CH₃ abundance would result in a lower concentration of CO in the stratosphere (Samuelson *et al.* 1983). On the other hand, CO may not be in a steady state and instead of being photochemically produced in the upper stratosphere it may be continuously replenished from the surface. The main difficulty encountered in explaining the CO mixing ratio profile of Fig. 2 is how to maintain a vertical gradient in the presence of even

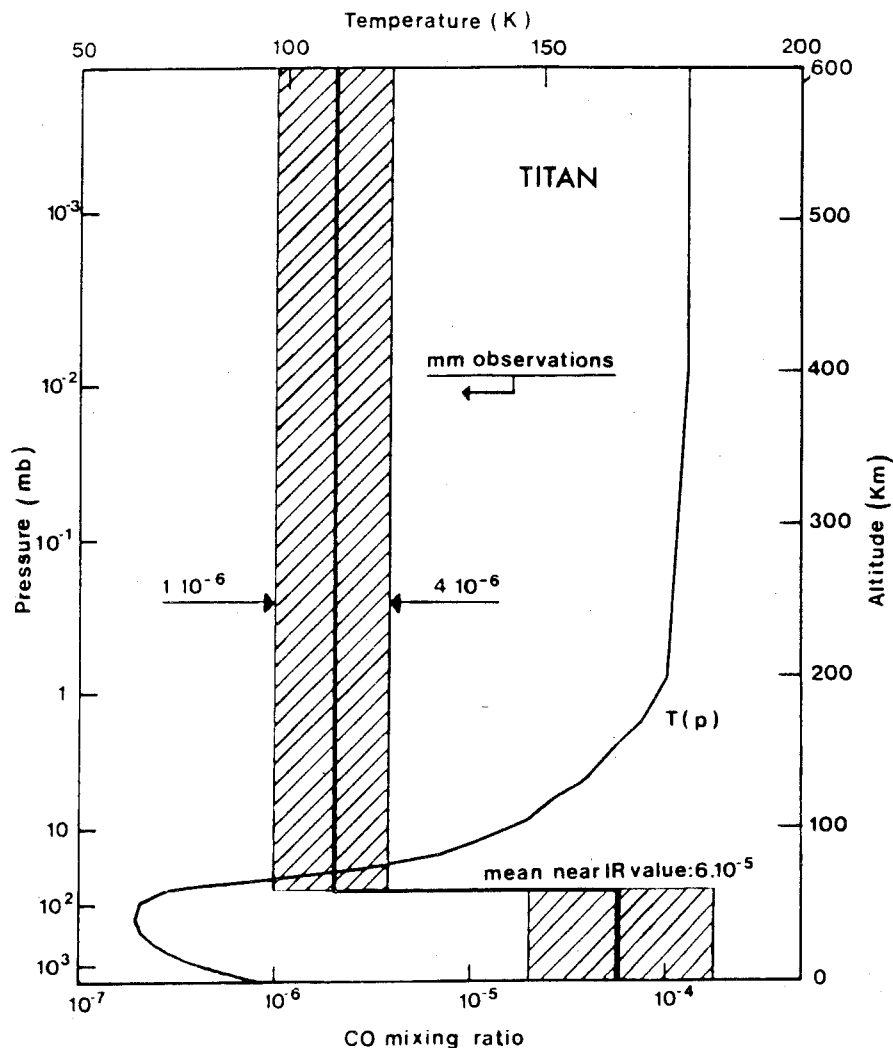


FIG. 2. "Error boxes" of vertical distributions of carbon monoxide in the Titan atmosphere. The tropospheric part is deduced from near-infrared measurements (Lutz *et al.* 1983). The stratospheric part corresponds to values derived from the present work. The thick line in the middle of the error boxes indicates the CO distribution which leads to the best fit of the feature located at the frequency of the (1-0) line on the observed spectrum (Fig. 1). Also displayed is the thermal profile $T(p)$ from Coustenis *et al.* (1987) used for calculating synthetic spectra.

minimal vertical mixing when current photochemical lifetimes predicted for CO in the Titan atmosphere are longer than 10^8 years. Rapid destruction of CO through the reaction $\text{CO} + \text{OH} \rightarrow \text{CO}_2 + \text{H}$ would require an amount of OH in the stratosphere several orders of magnitude higher than predicted by models.

If a CO gradient really exists near the tropopause level, the CO stratospheric abundance could also vary with latitude on the Titan disk. It is unfortunately impossible to look for such a variation with ground-based millimeter telescopes. It would require measurement of the CO rotation lines either in the far-infrared or in the millimeter range from a spacecraft orbiting Titan, as planned by the ESA-NASA Cassini Project.

ACKNOWLEDGMENTS

We are grateful to Dennis Downes and the staff of IRAM for their very efficient assistance. We thank

Eric Gérard for his help in analyzing the data, and Bruno Bézard and Régis Courtin for their comments on the manuscript. Stimulating discussions were held with Duane Muhleman. We thank one of our referees (T. Clancy) for his constructive comments, and especially for pointing out to us the true nature of the problem resulting from a vertical gradient in the CO mixing ratio. This work was supported by a grant from the ATP Planétologie de l'Institut National des Sciences de l'Univers (France).

REFERENCES

- BAARS, J. W. M., B. G. HOOGHOUT, P. G. MEZGER, AND M. J. DE JONGE 1987. The IRAM 30-m millimeter radio telescope on Pico-Veleta, Spain. *Astron. Astrophys.* **175**, 319-326.
- COLMONT, J. M., AND N. MONNANTEUIL 1986. Self, nitrogen and oxygen broadening of the 115 GHz line

- of carbon monoxide. *J. Quant. Spectrosc. Radiat. Transfer* **35**, 81-85.
- COURTIN, R. 1982. The spectrum of Titan in the far infrared and microwave regions. *Icarus* **51**, 466-475.
- COURTIN, R. 1988. Pressure-induced absorption coefficients for radiative transfer calculations in Titan's atmosphere. *Icarus*, in press.
- COUSTENIS, A., B. BÉZARD, D. GAUTIER, A. MARTEN, AND R. SAMUELSON 1987. Vertical distribution of hydrocarbons and nitriles in the Titan atmosphere. *Bull. Amer. Astron. Soc.* **19**, 873.
- DE BERGH, C., B. L. LUTZ, AND T. OWEN 1986. The troposphere of Titan probed by ground-based spectroscopy at 1.6 μm . *Bull. Amer. Astron. Soc.* **18**, 808-809.
- GÉRARD, E. 1986. The OH radio line in comets: A review. In *Cometary Radio Astronomy Proceedings of a NRAO Workshop held at The National Radio Astronomy, Green Bank, West Virginia, Sept. 24-26* (W. M. Irvine, P. Schloerb, and L. E. Tacconi, Eds.). Garman, Workshop No. 17.
- GUILLOTEAU, S. 1987. *Spectral Line Calibration of the 30-m and 15-m Antennas*. IRAM internal report, March.
- LINDAL, G. F., G. E. WOOD, H. B. HOTZ, D. N. SWEETNAM, V. R. ESHLEMAN, AND G. L. TYLER 1983. The atmosphere of Titan: An analysis of the Voyager 1 radio-occultation measurements. *Icarus* **53**, 348-363.
- LUTZ, B. L., C. DE BERGH, AND T. OWEN 1983. Titan: Discovery of carbon monoxide in its atmosphere. *Science* **220**, 1374-1375.
- MAUERSBERGER, R., M. GUELIN, J. MARTIN-PINTADO, AND J. CERNICARO 1987. Line calibrators for the 30-m telescope at $\lambda = 1.3, 2$ and 3-mm. Preprint.
- MUHLEMAN, D. O., G. L. BERGE, AND R. T. CLANCY 1984. Microwave measurements of carbon monoxide on Titan. *Science* **220**, 1374-1375.
- PAUBERT, G., D. GAUTIER, AND R. COURTIN 1984. The millimeter spectrum of Titan: Detectability of HCN, HC_3N and CH_3CN and the CO abundance. *Icarus* **60**, 599-612.
- PAUBERT, G., A. MARTEN, C. ROSOLEN, D. GAUTIER, AND R. COURTIN 1987. First radio detection of HCN on Titan. *Bull. Amer. Astron. Soc.* **19**, 633.
- POYNTER, R. L., AND H. M. PICKETT 1982. *Submillimeter, Millimeter and Microwave Spectral Line Catalog*, JPL Publ. 80-23, rev. 1. Jet Propulsion Laboratory, Pasadena, CA.
- SAMUELSON, R. E., W. C. MAGUIRE, R. A. HANEL, V. G. KUNDE, D. E. JENNINGS, Y. L. YUNG, AND A. C. AIKIN 1983. CO_2 on Titan. *J. Geophys. Res.* **88**, 8709-8715.
- YUNG, Y. L., M. ALLEN, AND J. P. PINTO 1984. Photochemistry of the atmosphere of Titan: Comparison between model and observations. *Astrophys. J. Suppl. Ser.* **55**, 465-506.

Article III

Stratospheric Profile of HCN on Titan from Millimeter Observations

LOÏC TANGUY,* BRUNO BÉZARD,* ANDRÉ MARTEN,* DANIEL GAUTIER,*
ERIC GÉRARD,* GABRIEL PAUBERT,† AND ALAIN LECACHEUX*

*Observatoire de Paris, Section de Meudon, 92195 Meudon, France, and †Institut de Radioastronomie
Millimétrique (IRAM), Granada, Spain

Received July 17, 1989; revised November 1, 1989

Measurements of the (1-0) line of hydrogen cyanide at 88.6 GHz in the Titan atmosphere are reported. Synthetic spectra were fitted to the observations to derive the vertical distribution of HCN in the stratosphere. The observed line is significantly narrower than that computed for constant stratospheric mixing ratios, implying an increase in the HCN concentration with altitude. From a least-squares analysis taking into account measurement noise and calibration uncertainties, a mean mixing ratio scale height of 47^{+36}_{-11} km is derived for the 100- to 300-km region. The HCN abundance is found to be best constrained around the 170-km level where the inferred mixing ratio is $3.3^{+0.9}_{-0.8} \times 10^{-7}$. The results are consistent with recent analyses of Voyager infrared measurements. The inferred vertical concentration gradient is much steeper and the abundance in the lower stratosphere smaller than predicted by current photochemical models. Theoretical HCN profiles may, however, be brought into agreement with the present results by reducing the magnitude of the vertical eddy mixing assumed in the stratosphere. © 1990

Academic Press, Inc.

I. INTRODUCTION

The detection of hydrogen cyanide HCN in the atmosphere of Titan by the infrared experiment (IRIS) aboard Voyager 1 (Hanel *et al.* 1981) is of considerable interest, HCN being a key intermediate in the synthesis of the organic molecules that led to the formation of life on Earth. The detection of other more complex nitriles (HC_3N , C_2N_2) and of many hydrocarbons obviously implies the existence of a complex chemistry operating in Titan's atmosphere. This idea has motivated the development of photochemical models for the satellite's atmosphere by Yung *et al.* (1984) and Yung (1987) who calculated vertical atmospheric distributions of the minor species observed (and of HCN in particular). These theoretical profiles remain to be compared with observations.

The IRIS spectrometer has recorded limb spectra at grazing incidence over the north

pole of Titan which permit the retrieval of the HCN vertical distribution in the altitude range 120 to 250 km, by analysis of the emission in the ν_2 band at $14 \mu\text{m}$ (Coustenis *et al.* 1987, Coustenis 1989). Information is also available at other latitudes, especially in the equatorial region, from observations at low emission angles, but the relatively low spectral resolution of IRIS restrains the inference of the HCN abundance to a limited atmospheric range centered around the 110-km altitude level (Coustenis *et al.* 1989). High-spectral-resolution infrared measurements from space may become available in the future, allowing us, in principle, to obtain a more precise vertical distribution, with however a persistent drawback set by the possible occurrence of a departure from LTE in the upper atmosphere.

Ground-based measurements of millimeter HCN emission lines provide a complementary approach to this question and ac-

tually present several advantages with respect to infrared observations. First, the extremely high spectral resolution achievable with heterodyne techniques allows us, in theory, to obtain the real shape of the emission line, then providing information on the HCN abundance up to very high atmospheric levels (Paubert *et al.* 1984). Second, the probability of spontaneous emission being proportional to the cube of the frequency, no significant departure from LTE is expected in the whole atmospheric range accessible to measurements. Third, aerosols and dust particles are not expected to absorb or scatter radiation at millimeter wavelengths. On the other hand, the small angular size of the satellite and its proximity to Saturn render millimeter heterodyne observations hard to obtain, and require the use of very large radiotelescopes, low-noise temperature receivers, and good observational conditions. Such facilities have recently become available in Europe.

The (1-0) line of HCN at 88 GHz was detected for the first time on Titan in September 1986 (Paubert *et al.* 1987) with the 30-m IRAM radiotelescope at Pico-Veleta (Spain). A second set of observations were successfully obtained at Pico-Veleta in May 1987. The purpose of the present article is to analyze these data to determine the vertical distribution of HCN in Titan's atmosphere.

Observations with the IRAM radiotelescope are described in Section II. Section III is devoted to the key problem of calibration. The procedure for retrieving the HCN concentration profile is described in Section IV. Results and error analysis are presented in Section V, and the results are discussed in Section VI.

II. OBSERVATIONS

The HCN observations reported here were carried out on 1986 September 7-8 and 1987 May 5-6 at the IRAM 30-m telescope (Baars *et al.* 1987), when the elongation of Titan with respect to Saturn was maximum. Saturn's maximal elevation was 33.8 and 31.6° in September and May, respectively. We employed the IRAM 80- to 115-GHz SIS mixer in September and the 74- to 115-GHz Schottky mixer in May. Both were used in the double-sideband (DSB) mode. The observed HCN line was located in the lower sideband. The DSB receiver noise temperature was ~200°K in May and slightly less in September. The observational data are summarized in Table I.

The back end consisted of a filter bank of 512 1-MHz-wide adjacent filters. In September, the whole back end was used, thus providing a 512-MHz-wide spectrum. In May, the back end was split into two halves, each 256 MHz wide: one half was dedicated to observation of the HCN (1-0)

TABLE I
OBSERVATIONAL PARAMETERS

	1986 September 7, 8	1987 May 5, 6
Effective frequency		88.632 GHz
Receiver	SIS	Schottky
DSB receiver noise temperature		~200°K
DSB system temperature		900-1200°K
Intermediate frequency		3.932 GHz
Half-power beam width (HPBW)		27 arcsec
Observational procedure		Position switching
Filter back end	512 × 1 MHz	256 × 1 MHz
Integration time on Titan	90 min	85 min
Calibration sources	Jupiter, Saturn	Saturn, IRC+10216

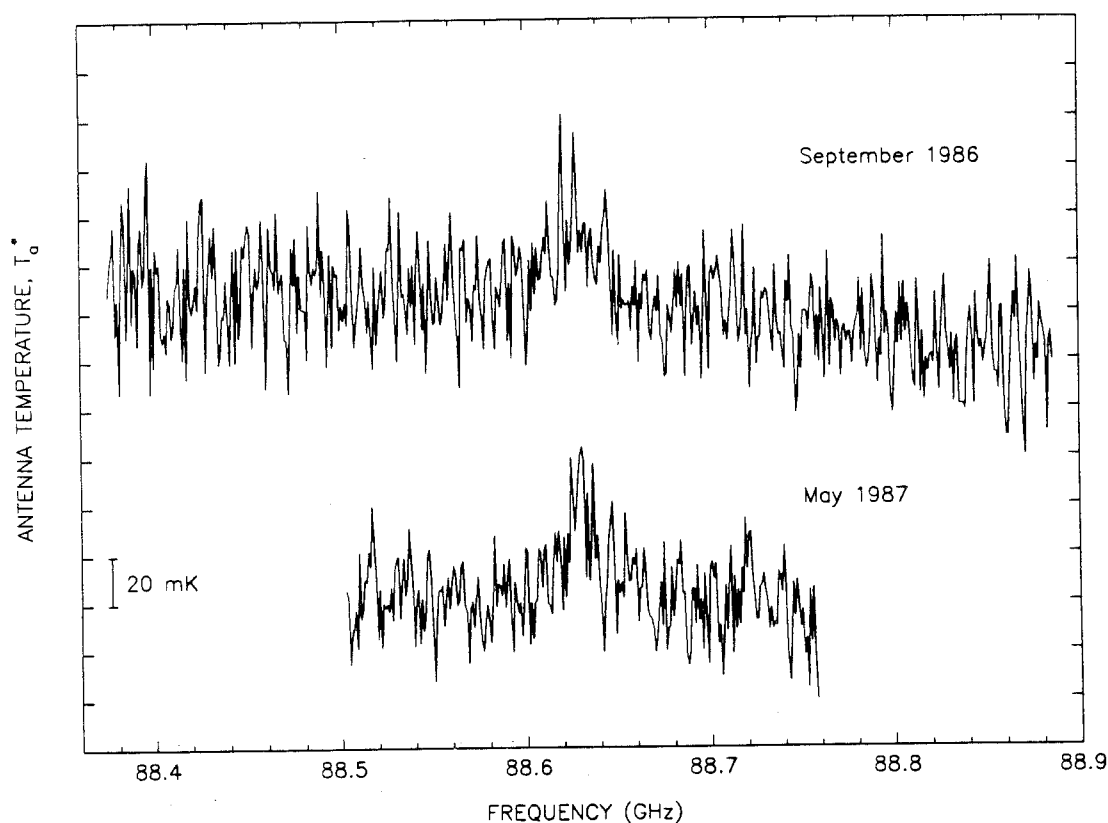


FIG. 1. Raw spectra of the $J = 0 \rightarrow 1$ line of HCN in the Titan atmosphere recorded in September 1986 and May 1987. The spectral resolution is 1 MHz. Data have been corrected for telluric opacity and antenna forward coupling efficiency. Continuum levels are arbitrary, and baselines have not been removed.

line at 88.6 GHz, and the other one to the CO (1-0) line at 115 GHz (Marten *et al.* 1988). The system temperature relative to outside the atmosphere, i.e., taking into account atmospheric absorption and emissivity and antenna efficiency, varied between 900 and 1200°K for both runs.

We used a position switching procedure which consisted of tracking alternatively 30 sec on Titan and 30 sec on a reference point off the source. The observed spectrum of Titan is then computed from the difference spectrum. To avoid an improbable contamination by Saturn coming through the antenna sidelobes (the elongation of Titan was about 3 arcmin, i.e., six times the antenna beamwidth), we chose a reference position symmetric to Titan with respect to Saturn's meridian. We performed 180 position switches in September and 170 in May. The actual integration times on Titan were thus

90 and 85 min, respectively. The corresponding spectra are shown in Fig. 1.

III. CALIBRATION

Three procedures were used to convert the raw data contrast into brightness temperature averaged over the Titan disk. The first two are the standard ones for line calibration at the 30-m telescope (Guilloteau 1987, Mauersberger *et al.* 1989), whereas the third one is specifically ours.

1. The first procedure (the skydip procedure) consists of measuring the total power on the sky at various elevations, and next on warm and cold loads. Bougher's law is adjusted to the data in a least-squares sense and provides the antenna forward efficiency F_{eff} (i.e., the antenna coupling to the sky in 2π steradians) and the zenith atmospheric opacity, which varies from 0.01 to 0.1 depending on weather conditions.

2. The second procedure (the cal cold procedure) computes the scaling factor from back-end counts (the raw data) to antenna temperatures above the atmosphere. It also provides the system and receiver temperatures. As for skydip, the receiver faces successively warm and cold matched loads. This procedure was repeated approximately every 20 min.

The two steps described above are usually sufficient for astronomical purposes. They convert the raw data into mean brightness temperature outside of the atmosphere, averaged over the whole visible sky (T_a^*). The forward efficiency F_{eff} is easy to measure and independent of the radiometer setup (at 90 GHz, $F_{\text{eff}} \approx 0.84$, i.e., all the rearward losses amount to 16%). Both procedures are entirely software supported on the telescope site, allowing us to directly compare our line calibration spectra (for example IRC+10216) with those of the IRAM catalog (Mauersberger *et al.* 1989) and to ascertain that our temperature scale was correct.

3. The third step consists of converting a brightness temperature averaged over the whole forward sky to a brightness temperature averaged over Titan's disk, whose radius is taken as 3175 km [this value corresponds to an altitude of 600 km above the surface, above which the gaseous opacity is negligible (Paubert *et al.* 1984)]. If the an-

tenna radiation pattern including the far sidelobes were accurately measured, computation of the so-called "dilution factor" to obtain the mean brightness temperature of Titan would be straightforward. In practice, the exact beam shape is not well known and only the first sidelobes are observed. Thus we have used measurements relative to different planets (Jupiter, Saturn) to evaluate the dilution factor for Titan. In this case, it is sufficient to know the main beam pattern, provided that one relies on published absolute brightness temperatures T_b .

The beam shape that we observed at 89 GHz from azimuth and elevation drift scans of point sources is not an Airy pattern (the illumination is tapered, reducing the sidelobe level); it is well fitted with a Gaussian (a) and an even better match would be provided by an Airy pattern truncated at the first null of the Bessel function (b). Adopting mean disk brightness temperatures of 180°K for Jupiter (de Pater and Massie 1985) and 150°K for Saturn (Epstein *et al.* 1984), we obtained the effective dilution factors (D_{obs}) defined as T_b/T_a^* and listed in Table II. Dilution factors can also be calculated if one neglects the contribution of the far sidelobes. A disk of uniform brightness temperature and given angular size is then convolved with beam patterns (a,b), yielding the results given in Table III (D_{calc}).

TABLE II
OBSERVED DILUTION FACTORS FOR PLANETS

	September 1986		May 1987
	Saturn	Jupiter	Saturn
Measured effective antenna temperature, T_a^*	23.5 ± 4^a	114^{+15a}_{-30}	26 ± 3^a
Mean disk brightness temperature, T_b	150^b	180^c	150^b
Dilution factor, T_b/T_a^*	$6.4^{+1.3}_{-1.0}$	$1.6^{+0.3}_{-0.2}$	$5.8^{+0.7}_{-0.6}$

^a The quoted error bars correspond to the intrinsic scatter of the measurements.

^b From de Pater and Massie (1985).

^c From Epstein *et al.* (1984).

TABLE III
CALCULATED DILUTION FACTORS^a

	September 1986			May 1987	
	Saturn	Jupiter	Titan	Saturn	Titan
(a) Gaussian beam	4.5	1.11	1520	3.8	1220
(b) Truncated Airy beam	4.0	1.03	1320	3.3	1060

^a Assuming no contribution from the far sidelobes.

The ratios of calculated (D_{calc}) to observed (D_{obs}) dilution factors of Jupiter and Saturn are $0.68 (\pm 0.06)$ and $0.60 (\pm 0.05)$ for the Gaussian and truncated Airy patterns, respectively. The ratios are smaller than unity because of the contribution of the far sidelobes to the forward radiation pattern, which must also be taken into account in the analysis of Titan's measurements. Thus, a simple rule of proportionality yields the effective dilution factors for Titan, as listed in Table IV: more precisely, the values previously calculated for Titan were divided by the ratio $D_{\text{calc}}/D_{\text{obs}}$ that we determined from planetary measurements. We adopted mean values of 2120 ± 240 and 1870 ± 220 for September and May, respectively. The errors quoted above correspond to the intrinsic scatter of the individual measurements relative to Jupiter and Saturn. But one should add the uncertainty on the determination of T_a^* itself and on the absolute brightness temperatures of Saturn and Jupiter. We have estimated that the two latter uncertainties bring 10% error

each and that all three errors may be added quadratically since their causes were independent. Consequently the cumulative relative error is 20% and will be considered a conservative limit. The dilution factors are thus taken as 2120 ± 420 and 1870 ± 370 for September 1986 and May 1987, respectively.

IV. METHOD OF ANALYSIS

The retrieval of the HCN vertical distribution from the two sets of observations is based on a direct comparison of the data with synthetic spectra generated by a radiative transfer program (Paubert *et al.* 1984). Calculations take into account the sphericity of Titan's atmosphere, which cannot be neglected considering the size of the atmosphere with respect to the satellite's core ($R = 2575$ km), as well as the decrease in gravitational acceleration with altitude. The radiative transfer equation is integrated up to an altitude $z_t = 600$ km above the surface. The derived brightness temperature (then

TABLE IV
DILUTION FACTORS ADOPTED FOR TITAN'S OBSERVATIONS^a

	September 1986		May 1987
	Saturn	Jupiter	Saturn
(a) Gaussian beam	2150	2160	1880
(b) Truncated Airy beam	2120	2020	1860
Adopted values ^b	2120 ± 420		1870 ± 370

^a Values are deduced from measurements on the planets.

^b See text for discussion of the uncertainties.

corresponding to an average over a disk of radius $R + z_t = 3175$ km) is converted into antenna temperature after division by the appropriate "dilution factor" given in Section III.

We employed the nominal temperature profile of Coustenis *et al.* (1989) derived from a simultaneous analysis of Voyager 1 radio-occultation data (Lindal *et al.* 1983) and infrared measurements in the ν_4 band of methane at $7.7 \mu\text{m}$ for Titan's equatorial region. Spectroscopic parameters for HCN were taken from the JPL catalog (Poynter and Pickett 1982). We used a Voigt lineshape with a collisional half-width of $0.177 (T/300)^{-0.80} \text{ cm}^{-1} \text{ atm}^{-1}$ (Rohart *et al.* 1987). The continuum opacity due to the $\text{N}_2\text{-N}_2$ pressure-induced absorption was modeled according to Courtin (1982). The so-derived absorption coefficients were found to be in good agreement with Courtin's (1988) calculations based on more recent laboratory measurements.

A precise comparison of the observations with synthetic spectra must cope with two limitations: (1) the zero-flux level in the data is unknown, (2) baseline ripples may slightly affect the shape of the observed spectra. The usual procedure for removing baseline ripples, which consists of subtracting a polynomial constrained to fit continuum data points outside of the line itself, cannot be applied to our observations. The difficulty comes from the large extent of the wings of the HCN emission line (several hundreds of megahertz), which maintains the flux above the $\text{N}_2\text{-N}_2$ continuum level even at the extremities of our bandpass (especially when a 256-MHz filter bank is used as was the case in May 1987). We have then settled on the following procedure: *for each calculated spectrum*, we first determine the second-degree polynomial which yields the best fitting of the difference between the raw observed spectrum and the synthetic one. Optimum values for the polynomial coefficients are obtained by minimizing χ^2 (Bevington 1969). The so-derived value of

χ^2 is then used as a measure of the quality of the agreement with the observations for the corresponding HCN vertical distribution.

We first performed radiative transfer calculations assuming a constant HCN mixing ratio in the stratosphere down to the condensation level. Following the comparison procedure described above, we find that the observations are best fitted for a mole fraction of 4×10^{-7} . Figure 2 shows the comparison of the corresponding synthetic spectrum (b) with the May data. The synthetic line does not indeed correctly reproduce the shape of the observed one which is significantly narrower. As illustrated in Fig. 2, use of a smaller (a) or larger (c) HCN abundance does not help to improve the fit, as the contrast between the line center and the wings (some 50 MHz away) is not increased. The discrepancy is actually suggestive of an altitude-dependent HCN mole fraction, as can be understood from Fig. 3. Contribution functions, calculated for a mixing ratio of 4×10^{-7} , are displayed at a resolution of 1 MHz for the line center at 88.632 GHz, as well as for 10 and 50 MHz further. They are defined as $(1/(B_\nu(T(p))/I_\nu)) d e^{-\tau_\nu}/d Lnp$, where I_ν is the outgoing spectral radiance at frequency ν , B_ν is the Planck function at temperature T , and τ_ν is the slant path optical depth at pressure level p , so that they indicate the relative contribution of each layer to the observed emission. Titan's flux originates partly in the troposphere where the $\text{N}_2\text{-N}_2$ opacity is large and partly in the stratosphere through the HCN line opacity. The stratospheric component, which is responsible for the observed emission above the continuum level, peaks around 0.2 mbar (270 km) at the line center and in the region 2–10 mbar (100–160 km) beyond 10 MHz. The narrowness of the observed emission feature (Fig. 2) indicates a decrease in the absorption coefficient in the line wings more pronounced than that calculated with a uniform HCN profile. This behavior can, in turn, be ex-

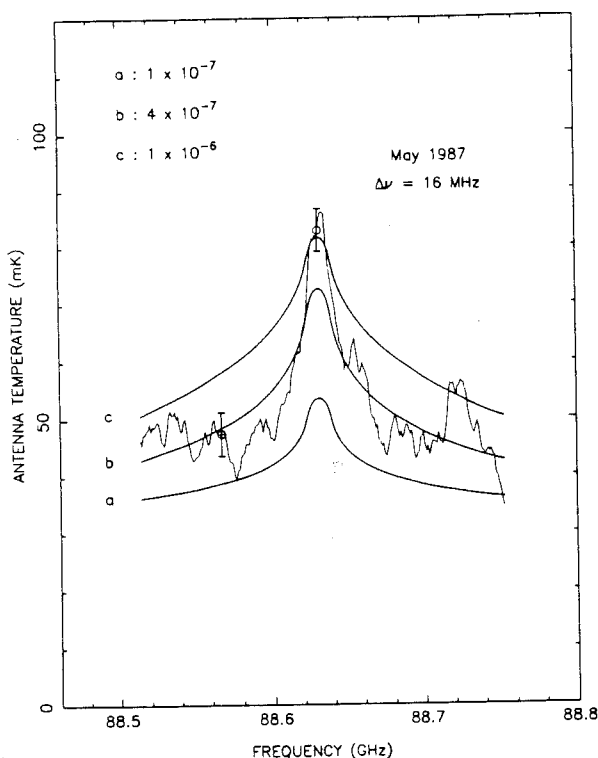


FIG. 2. Comparison of the May observations with synthetic spectra including HCN mole fractions of 1×10^{-7} (a), 4×10^{-7} (b) and 1×10^{-6} (c). Spectra have been downgraded to a resolution of 16 MHz to enhance visibility. Data points have been fitted in a least-squares sense to spectrum (b), following the procedure described in the text. The bars indicate the 1SD noise level for the 16-MHz resolution spectrum.

plained by a decrease in the HCN mixing ratio with depth, since the emission originates from significantly deeper levels in the wings than in the line center.

To investigate the altitude dependence of the HCN concentration, we have considered two-parameter vertical distributions in the form of $q(z) = q(z_0)e^{(z-z_0)/H}$, where $q(z_0)$ is the mole fraction at altitude $z_0 = 100$ km and H is a vertical mixing ratio scale height. An upper boundary condition is set where the concentration is held constant above the 425-km level to avoid unrealistically high values in the uppermost atmospheric layers. The HCN distribution is further constrained to follow the saturation law in the lower atmosphere, below the condensa-

tion level which is located around the 80-km level. Synthetic spectra were then generated for a set of 7×7 HCN vertical profiles defined by the couple of parameters $(q(z_0), 1/H)$, with $q(z_0)$ varying from 0.1 to 4×10^{-7} and $1/H$ from 0 to 0.04 km^{-1} . Values of χ^2 were next calculated for each distribution and for both the September and the May spectra, following the procedure described above. Contour plots of χ^2 in the two-parameter space were finally obtained from spline interpolation of the initial 7×7 grid. The best-fitting HCN distribution results from the location of the minimum value for χ^2 (χ^2_{\min}) while the surface where $\chi^2 \geq \chi^2_{\min} + n^2$ defines a range of acceptable solutions within a $n\sigma$ confidence level.

V. RESULTS AND ERROR ANALYSIS

Results of the analysis successively applied to the September and May data sets are presented in Fig. 4a. The best-fitting distributions are characterized by $q(100 \text{ km}) = 1.5 \pm 0.8 \times 10^{-7}$ and $H = 62^{+28}_{-23} \text{ km}$ for September observations, $q(100 \text{ km}) = 0.45^{+0.35}_{-0.25} \times 10^{-7}$ and $H = 42 \pm 10 \text{ km}$ for May observations. The error bars are related to a 1σ confidence level. The residuals between the calculated and observed spectra then reach their minimal values, 17.4 and 14.7°mK for September and May, respectively.

It is worth noting that the abscissa axis, which corresponds to uniformly mixed HCN distributions, does not intercept the 2σ contour lines relative to either September or May observations. A mixing ratio of 4×10^{-7} , which minimizes the residuals of fit when a constant concentration profile is used, actually leads to increases in χ^2 of 7 (for September) and 17 (for May) from its minimum value χ^2_{\min} . From the analysis of the May spectra, altitude-constant profiles can thus be excluded at an $\sim 4\sigma$ confidence level, which confirms the visual impression of Fig. 2. Although the S/N ratio is lower in the September measurements, uniformly

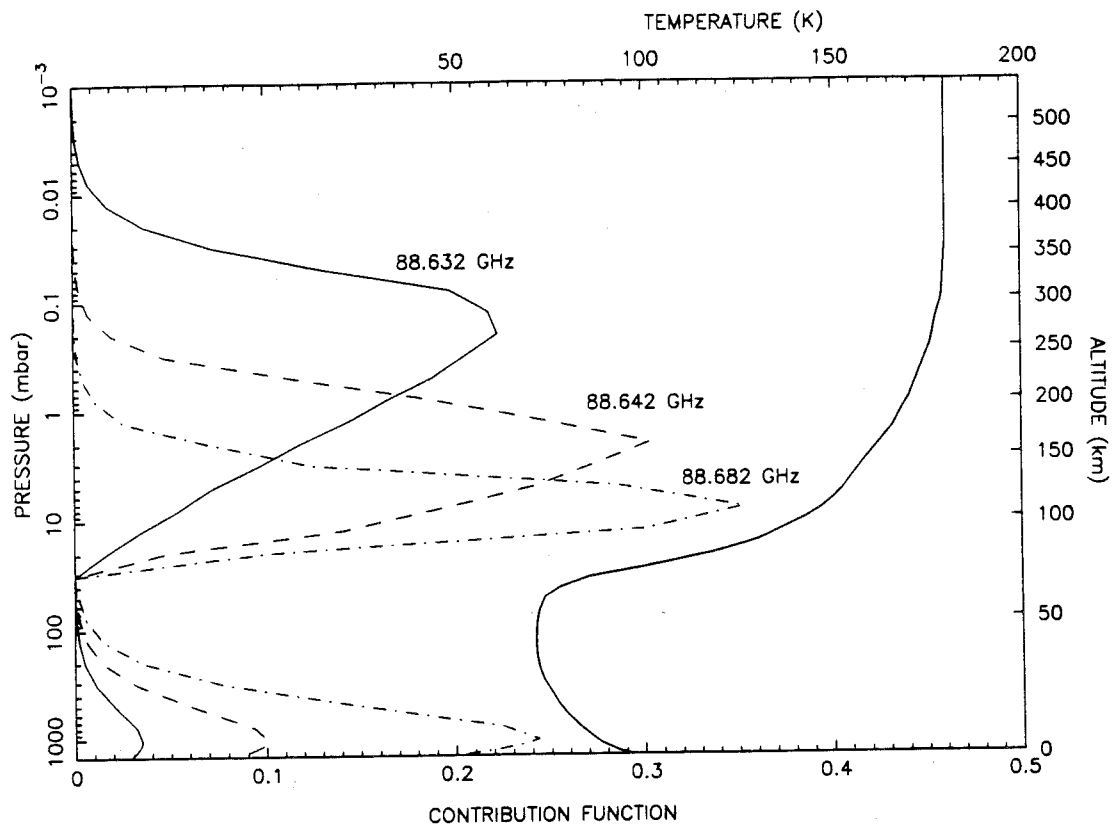


FIG. 3. Normalized contribution functions at a resolution of 1 MHz for line center (88.632 GHz) and in the line wings at 88.642 and 88.682 GHz, indicating the altitude range where the outgoing emission originates. Also shown is the temperature profile used in the computations.

mixed HCN profiles are still at odds with the data at the 2.7σ level.

We consider that the analyses of the two data sets are not at variance, since the 1σ surfaces in the two-parameter space are almost tangential to one another while the optimum distribution associated with one set of observations essentially lies on the 2σ contour line corresponding to the other one. In a second step, we have thus determined the vertical distribution that yields the best agreement (in a least-squares sense) with the two observations simultaneously, with appropriate weighting accounting for the difference in the rms noise of the data. Results are presented in Fig. 4b. Parameters for the best-fitting HCN profile are then $q(100 \text{ km}) = 0.75^{+0.45}_{-0.25} \times 10^{-7}$ and $H = 47^{+12}_{-6} \text{ km}$, where the quoted uncertainties refer to a 1σ confidence level. The spectrum generated with this distribution is compared to the September and May measurements in Fig. 5. The HCN distribu-

tions, with parameters located on the 1σ contour line of Fig. 4b, define a family of profiles $\{q(z)\}$ bounded by two envelopes which provide an estimate of the level-by-level uncertainty in the HCN mixing ratio (column 2 of Table V).

Besides random noise, systematic errors may arise from the data calibration procedure described in Section II. As discussed there, a conservative value of $\pm 20\%$ can be assumed for the accuracy of the temperature scale. To estimate how this uncertainty is propagated into the derived HCN mixing profile, we have successively multiplied the observed spectra by 1.2 and 0.8 and applied the same analysis as for the nominal data. The vertical distributions that yield the best fitting of the so-altered data are displayed in Fig. 6, along with the nominal HCN profile. These two extreme profiles can be regarded as boundaries defining the uncertainty zone for the HCN vertical distribution insofar as calibration errors solely are concerned. The

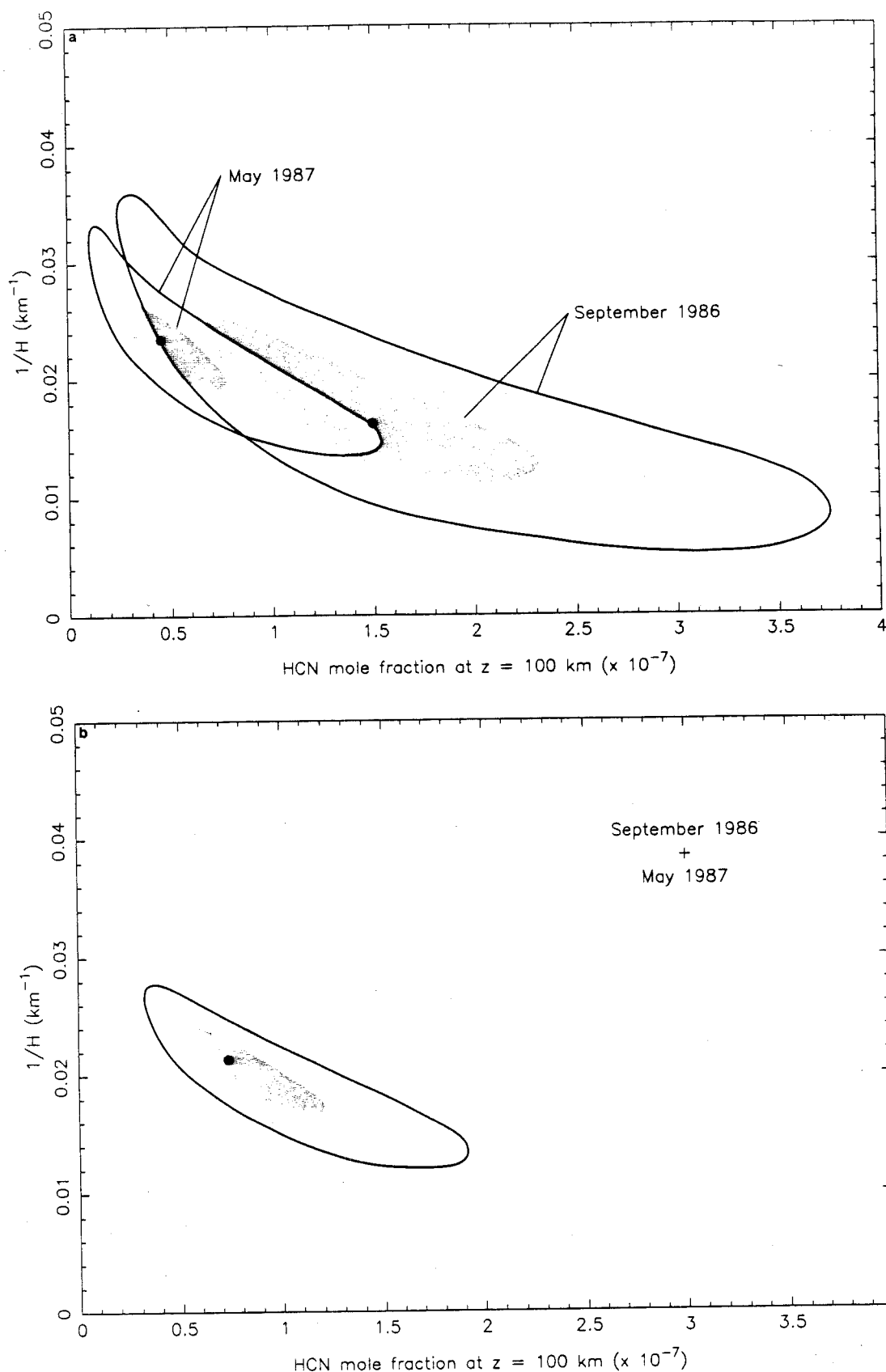


FIG. 4. Results of radiative transfer model calculations. The shaded regions show the model parameters which agree with the data within a 1σ confidence level. Solid lines indicate 2σ contours. The two model parameters are the HCN mole fraction at 100 km and the inverse of the HCN mixing ratio scale height. Note that the abscissa axis ($1/H = 0$) corresponds to uniformly mixed distributions above the condensation level. (a) Fitting of the May and September observations, considered as two independent data sets. (b) Simultaneous fitting of the two sets of observations.

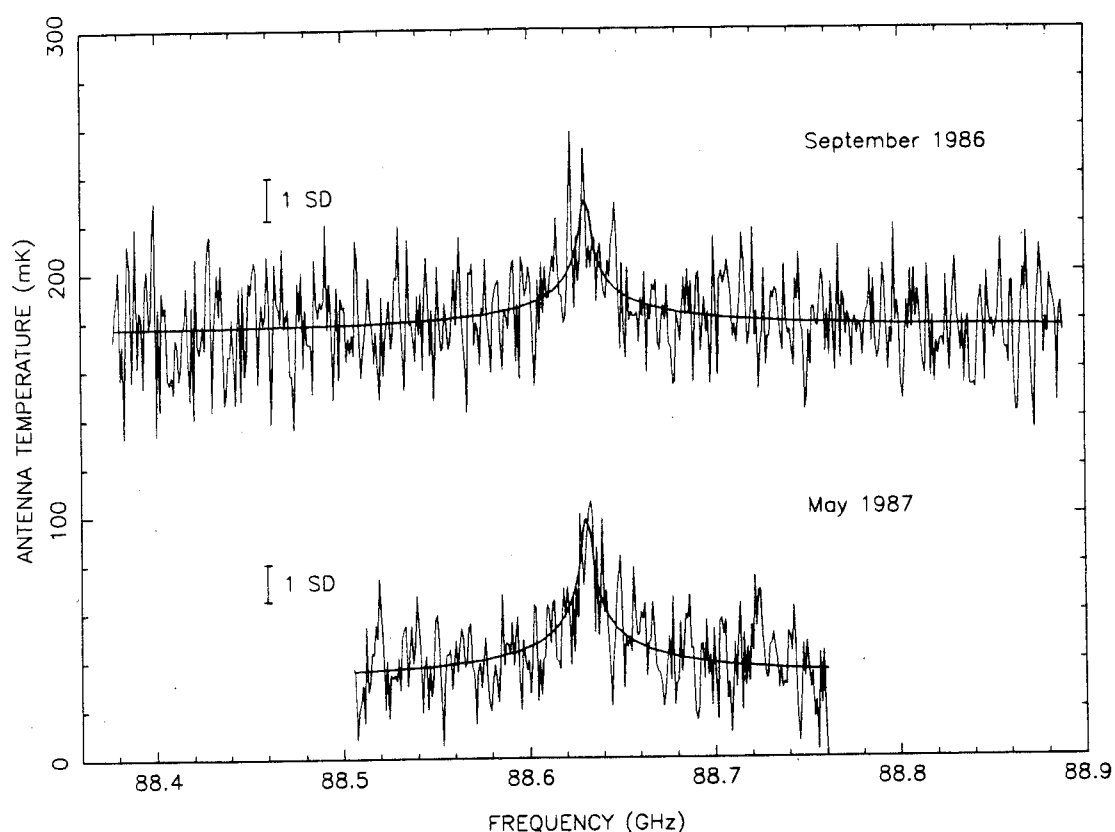


FIG. 5. Comparison of observations with the best-fit synthetic spectrum generated with the HCN vertical distribution shown in Figs. 6 and 7 (heavy line). Spectra relative to September observations have been shifted by 150°mK. The bars indicate the 1SD noise level per datum point for the September and May spectra.

corresponding error bars on the HCN concentration are listed for a few altitudes in Table V (column 3).

Uncertainties on the model atmosphere used in the calculation of synthetic spectra

must also be considered. Spectroscopic data are expected to be accurate to within a few percent. With respect to the temperature profile, Lellouch *et al.* (1989) and Coustenis *et al.* (1989) have recently reana-

TABLE V
RESULTS OF THE VERTICAL DISTRIBUTION OF HCN

	Best-fitting distribution	Uncertainty		
		Measurement noise (1SD)	Calibration (±20%)	Total
Mixing ratio scale height (km)	47	+12/-6	+34/-9	+36/-11
Mixing ratio ($\times 10^{-7}$)				
100 km	0.75	+0.45/-0.25	+0.50/-0.20	+0.8/-0.3
170 km	3.3	+0.9/-0.7	+0.15/-0.3	+0.9/-0.8
200 km	6.2	+1.2/-1.2	+1.4/-1.9	+1.9/-2.1
300 km	52.	+23/-20	+56/-37	+66/-39

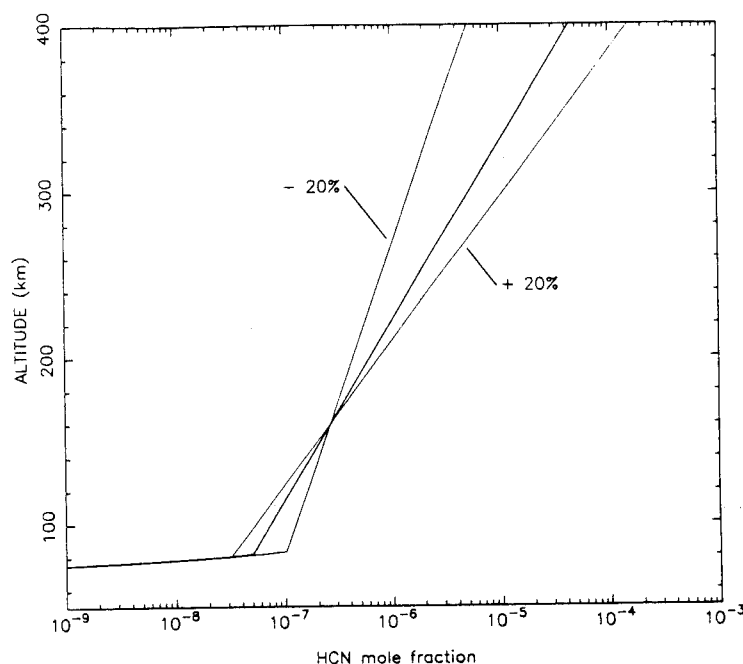


FIG. 6. Effect of a $\pm 20\%$ variation in the calibration factor applied simultaneously to the two sets of observations. Heavy line: the HCN vertical distribution yielding the best fit of the nominal spectra shown in Fig. 5; thin lines: vertical profiles derived from the fitting of antenna temperature spectra multiplied by 1.2 and 0.8.

lyzed the Voyager 1 radio-occultation data (Lindal *et al.* 1983), in conjunction with infrared observations in the $7.7\text{-}\mu\text{m}$ methane band. Coustenis *et al.* proposed (1) a nominal profile, corresponding to an argon-free atmosphere, which was used in the present study; (2) envelope profiles bounding a range of possible solutions, taking into account noise in the refractivity data and uncertainties in atmospheric composition. In fact, these two extreme profiles do not deviate from the nominal one by more than $\pm 5\%$ (see Fig. 5 of Coustenis *et al.* 1989). Because of the linear variation of the Planck function with temperature, the uncertainty thus imposed on the calculated spectra should not exceed $\pm 5\%$. This is much less than the $\pm 20\%$ that we retained for the calibration uncertainty, and we have accordingly neglected systematic errors coming from the atmospheric model.

We have then calculated, level by level, the quadrature sum of the two relative error terms coming from random noise and a $\pm 20\%$ calibration uncertainty (column 4 of Table V). This leads to the two extreme vertical distributions displayed in Fig. 7, which reflect the total uncertainty in the HCN mixing profile. These calculations pertain to the region of the atmosphere which significantly contributes to the HCN emission line, i.e., the 100- to 300-km range (see Fig. 3). The error analysis indicates that the uncertainty resulting from noise level prevails over that due to calibration between 100 and 190 km, while the latter predominates in the overlying atmosphere. The relative error in the HCN mixing ratio reaches its minimum ($\sim \pm 30\%$) at an altitude of 170 km, where $q(\text{HCN}) = 3.3^{+0.9}_{-0.8} \times 10^{-7}$. Hence, this is the level for which maximum information is available from our measurements.

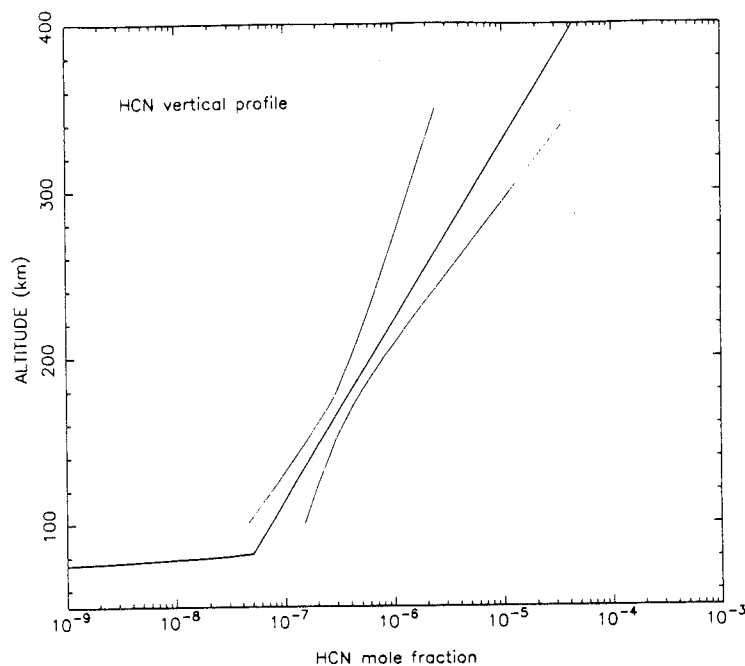


FIG. 7. Vertical distribution of HCN in the atmosphere of Titan, derived from the observed emission spectra in Fig. 1. Boundaries about the heavy-line best-fit profile indicate the uncertainty in the HCN mixing profile from both the measurement noise and a possible error of $\pm 20\%$ in the flux calibration.

VI. DISCUSSION

Our analysis shows that hydrogen cyanide is not uniformly mixed above its condensation level. A mole fraction of 4×10^{-7} , which minimizes the discrepancy with the observations when a uniformly mixed distribution is assumed, is still discordant with the data at the 5σ level. On the contrary, the observed spectra can be correctly reproduced assuming a pronounced increase in the HCN concentration with altitude. From a least-squares fitting analysis, we derive a mean mixing ratio scale height of 47^{+36}_{-11} km for the 100- to 300-km region. The HCN vertical distribution is best constrained by the measurements around the 170-km level for which a mixing ratio of $3.3^{+0.9}_{-0.8} \times 10^{-7}$ is inferred. These results refer to planetary averaged conditions since Titan's disk is not resolved.

On the other hand, spatial resolution is available from the Voyager IRIS measurements of Titan's thermal emission covering

the 200- to 1500-cm^{-1} range at a 4.3-cm^{-1} resolution. Coustenis *et al.* (1989) have analyzed selections of IRIS spectra recorded in the equatorial region of Titan, where the thermal structure is best defined, as a result of the combined analysis of radio-occultation data and infrared measurements in the $7.7\text{-}\mu\text{m}$ methane band. From the best fit of the ν_2 band of HCN centered at 713 cm^{-1} ($\lambda \sim 14\text{ }\mu\text{m}$), they derived a mean stratospheric mole fraction of $1.6^{+0.4}_{-0.6} \times 10^{-7}$. The limited spectral resolution precludes the retrieval of a vertical distribution and the inferred mixing ratio is in fact representative of an atmospheric region centered at the 110-km altitude level. The IRIS-derived abundance is compatible with our result for this altitude ($q = 0.9^{+0.8}_{-0.3} \times 10^{-7}$), taking into account the error bars attached to the two determinations. On the other hand, horizontal viewing observations acquired by Voyager 1 at Titan's north pole allow the retrieval of the local HCN vertical distribution. The HCN mole fraction steeply in-

creases with height, rising from 2.4×10^{-7} at the 120-km altitude level up to 1.6×10^{-6} at 250 km (Coustenis *et al.* 1987, 1989). The mean mixing ratio scale height in this altitude range is then ≈ 70 km, which is consistent with the vertical gradient we inferred considering the uncertainties. Comparison of the two results suggests that the stratospheric abundance of HCN rapidly increases with altitude everywhere on Titan's disk, a behavior which provides useful constraints on photochemical models as discussed below.

Yung *et al.* (1984) have developed a one-dimensional photochemical model of Titan's atmosphere which incorporates chemical schemes for numerous compounds bearing carbon, hydrogen, nitrogen, and oxygen atoms. Exospheric boundary conditions and assumptions on the vertical mixing rates were also included in the model. More recently, the photochemistry of nitrile compounds was updated by Yung (1987) in the light of new laboratory kinetic measurements.

According to these photochemical calculations, hydrogen cyanide is formed mainly in the thermosphere, above the 800-km level, from CH_3 and N radicals, the latter being derived from dissociation of N_2 driven by energetic electrons of magnetospheric origin. Small amounts of HCN are converted into more complex nitrile species (HC_3N , C_2N_2 , C_4N_2) in the lower atmosphere. Nitrile compounds are subsequently transported by diffusion down to observable levels in the stratosphere, and are ultimately lost by condensation in the ~ 80 -km region. The predicted altitude profile for HCN thus decreases with depth in the stratosphere, in agreement with our results. However, the vertical gradient that we derived is much stronger than in the Yung *et al.* (1984) and Yung (1987) models where the computed mixing ratio increases by less than 20% from 150 to 250 km. Furthermore, the predicted concentration in this altitude range is $\sim 10^{-5}$ which is excluded by our observations. On the other

hand, the model leads to a strong decrease in the HCN abundance below the 140-km level. The behavior of the computed altitude profile is linked to model assumptions on the eddy diffusion rates. Below 700 km, the HCN abundance is governed mainly by diffusion processes, while UV photodissociation and recycling through reactions of ethane with CN radicals should play a less important role (e.g., Fig. 9 of Yung *et al.* 1984). Yung *et al.* have parametrized vertical transport mechanisms with the so-called eddy diffusion coefficient K . They have chosen a somewhat arbitrary profile for $K(z)$ and have adjusted its magnitude in the thermosphere ($z \geq 700$ km) and in the lower stratosphere ($z \leq 140$ km) so that the model yields column-averaged mixing ratios for various species similar to the first published results derived from Voyager observations. In the framework of this formalism, the quantity $Kndq/dz$, where n is the atmospheric number density, must remain constant with altitude z , provided that destruction and recycling reactions for HCN can be neglected, as is roughly the case below the ~ 700 -km level. The mixing ratio gradient is then inversely proportional to K in the stratosphere, so that the mixing ratio itself q is extremely sensitive to the choice of the eddy diffusivity profile. Hence, the discrepancy between the photochemical model results and our measurements does not necessarily imply that the important chemical schemes for nitrile production have not been correctly identified, but could be related to uncertainties in the vertical mixing rates.

To further investigate this idea, we have tried to satisfy the constraints brought by the microwave observations, by modifying in the stratosphere the eddy mixing profile $K(z)$ used by Yung *et al.* We first assume that the model input and output parameters are valid above 700 km and we use the values of $K(z)$ and $q(z)$ at this level as upper boundary conditions for our calculation. The altitude-dependent mixing ratio $q(z)$ is obtained from downward integration of the

equation

$$Kndq/dz = C, \quad (1)$$

where C is a constant which is actually equal to the integrated photochemical production rate of HCN in the thermosphere. We have determined its value ($\sim 10^8 \text{ cm}^{-2} \text{ sec}^{-1}$) from the HCN mixing profile obtained by Yung *et al.* (Fig. 8 of that article), and we assume that any change in $K(z)$ in the lower stratosphere will not affect it significantly. Inserting the observational constraint on the HCN gradient in the 170-km region ($H = 47_{-11}^{+36} \text{ km}$) into Eq. (1), we find that $K(170 \text{ km})$ must be located between 2.5 and $6 \times 10^4 \text{ cm}^2 \text{ sec}^{-1}$, with a preferred value of $3.4 \times 10^4 \text{ cm}^2 \text{ sec}^{-1}$. Figure 8 shows two possible vertical profiles for $K(z)$ which satisfy this constraint. In the first

case (dashed line), K is taken to be constant up to a cutoff altitude above which the Yung *et al.* profile is used. This somewhat artificial discontinuity simulates the transition to a region of stagnancy expected to occur in the lower stratosphere due to thermal inversion. In the second case (dotted line), K increases with altitude according to a constant scale height up to the level where it joins the original curve of Yung *et al.* The disconnection levels are determined so that the corresponding HCN altitude profiles agree with the observational constraints. As can be seen in the left side of Fig. 8, the so-derived HCN distributions do not depart from the allowed uncertainty zone (shaded area) in the whole 100- to 300-km range. An infinity of other eddy mixing profiles would obviously be capable of

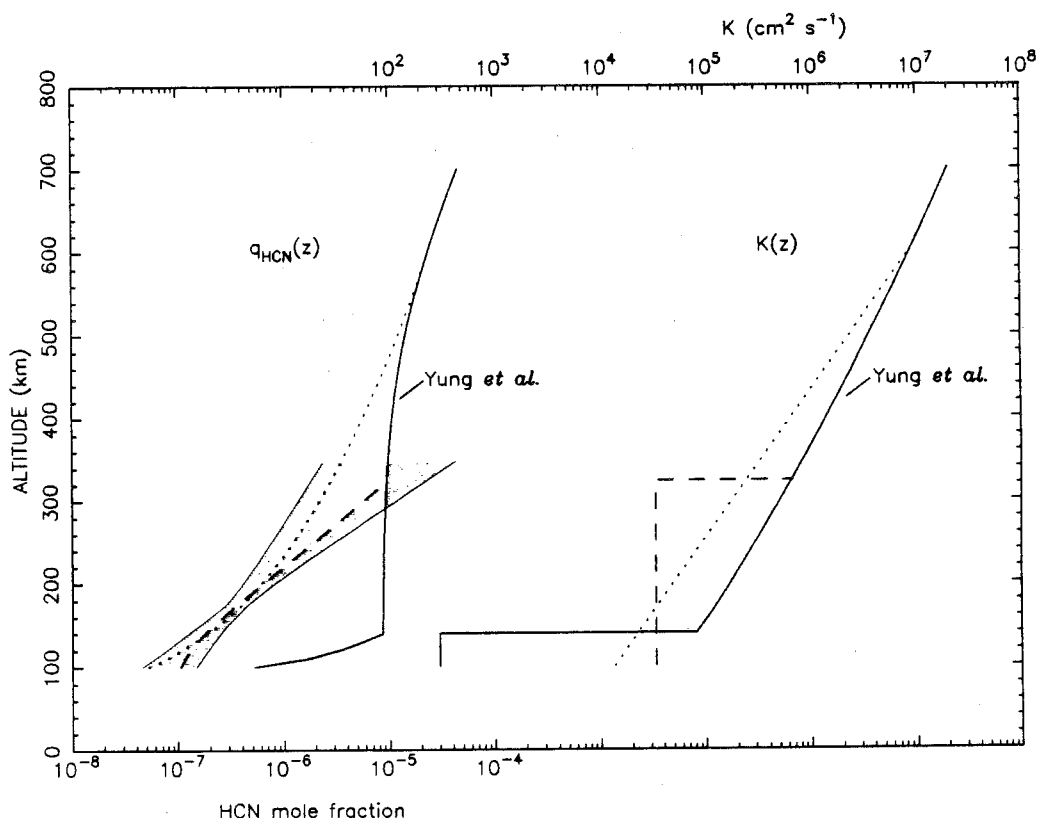


FIG. 8. Vertical distributions of HCN $q_{\text{HCN}}(z)$ calculated for various altitude profiles of the eddy diffusion coefficient $K(z)$. Predictions from the photochemical model of Yung *et al.* (1984) are used as boundary conditions above the 700-km level. Solid lines: the Yung *et al.* model; dashed lines: $K(z) = 3.4 \times 10^4 \text{ cm}^2 \text{ sec}^{-1}$ for $z < 324 \text{ km}$; dotted lines: $K(z)$ varies as $e^{z/H}$ ($H = 77.4 \text{ km}$) for $z < 600 \text{ km}$. The shaded area defines the range of acceptable HCN profiles as constrained by the millimeter observations.

yielding acceptable HCN concentrations, provided that, at least below 320 km, they diverge from the profile assumed by Yung *et al.* and that K reaches $2.5\text{--}6 \times 10^4 \text{ cm}^2 \text{ sec}^{-1}$ around 170 km.

The above calculations rely on the Yung *et al.* model results for the $z \geq 700\text{-km}$ region [as concerns $K(z)$ and $q(z)$]. Note that, according to the recent work of Yung (1987), the constant C in Eq. (1) (determined from Fig. 1 or 3 of that paper) works out at a two to three times higher value so that $K(170 \text{ km})$ should be likewise increased. With these caveats in mind, we can nevertheless conclude that to reconcile photochemical calculations with the present millimeter measurements, it is necessary to consider eddy diffusion coefficients smaller than those of Yung *et al.* (1984), at least over some region of the stratosphere below $\sim 320 \text{ km}$.

ACKNOWLEDGMENTS

We are grateful to the IRAM staff members for their efficient assistance during the observations. We thank A. Coustenis for providing helpful comments on the manuscript. This work was supported by a grant from the ATP-Planétologie de l'Institut National des Sciences de l'Univers.

REFERENCES

- BAARS, J. W. M., B. G. HOOGHOUTD, P. G. MEZGER, AND M. J. DE JONGE 1987. The IRAM 30-m millimeter radio telescope on Pico-Veleta, Spain. *Astron. Astrophys.* **175**, 319–326.
- BEVINGTON, P. R. 1969. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, New York.
- COURTIN, R. 1982. The spectrum of Titan in the far-infrared and microwave regions. *Icarus* **51**, 466–475.
- COURTIN, R. 1988. Pressure-induced absorption coefficients for radiative transfer calculations in Titan's atmosphere. *Icarus* **75**, 245–254.
- COUSTENIS, A. 1989. *L'atmosphère de Titan à partir des observations infrarouges de Voyager*. Thèse de doctorat, Université de Paris 7.
- COUSTENIS, A., B. BÉZARD, AND D. GAUTIER 1989. Titan's atmosphere from Voyager infrared observations. I. The gas composition of Titan's equatorial region. *Icarus* **80**, 54–76.
- COUSTENIS, A., B. BÉZARD, D. GAUTIER, AND A. MARTEN. 1987. Vertical distributions of hydrocarbons and nitriles in the Titan atmosphere. *Bull. Amer. Astron. Soc.* **19**, 873.
- DE PATER, I., AND S. T. MASSIE 1985. Models of the millimeter-centimeter spectra of the giant planets. *Icarus* **62**, 143–171.
- EPSTEIN, E. E., M. A. JANSSEN, AND J. N. CUZZI 1984. Saturn's rings: 3-mm low-inclination observations and derived properties. *Icarus* **58**, 403–411.
- GUILLOTEAU, S. 1987. *Spectral Line Calibration of the 30-m and 15-m Antennas*. IRAM report.
- HANEL, R., B. CONRATH, F. M. FLASAR, V. KUNDE, W. MAGUIRE, J. PEARL, J. PIRAGLIA, R. SAMUELSON, L. HERATH, M. ALLISON, D. CRUIKSHANK, D. GAUTIER, P. GIERASCH, L. HORN, R. KOPPANY, AND C. PONNAMPERUMA 1981. Infrared observations of the Saturnian system from Voyager 1. *Science* **212**, 192–200.
- LELLOUCH, E., A. COUSTENIS, D. GAUTIER, F. RAULIN, N. DUBOULOZ, AND C. FRÈRE 1989. Titan's atmosphere and hypothesized ocean: A reanalysis of the Voyager 1 radio-occultation and IRIS $7.7 \mu\text{m}$ data. *Icarus* **79**, 328–349.
- LINDAL, G. F., G. E. WOOD, H. B. HOTZ, D. N. SWEETNAM, V. R. ESHLEMAN, AND G. L. TYLER 1983. The atmosphere of Titan: An analysis of the Voyager 1 radio occultation measurements. *Icarus* **53**, 348–363.
- MARTEN, A., D. GAUTIER, L. TANGUY, A. LECAHEUX, C. ROSOLEN, AND G. PAUBERT 1988. Abundance of carbon monoxide in the stratosphere of Titan from millimeter heterodyne observations. *Icarus* **76**, 558–562.
- MAUERSBERGER, R., M. GUÉLIN, J. MARTIN-PINTADO, J. CERNICARO, H. HEIN, AND S. NAVARRO 1989. Line calibrators at $\lambda = 1.3, 2$ and 3 mm . *Astron. Astrophys. Suppl. Ser.* **79**, 217–261.
- PAUBERT, G., D. GAUTIER, AND R. COURTIN 1984. The millimeter spectrum of Titan: Detectability of HCN, HC_3N , and CH_3CN and the CO abundance. *Icarus* **60**, 599–612.
- PAUBERT, G., A. MARTEN, C. ROSOLEN, D. GAUTIER, AND R. COURTIN 1987. First radio detection of HCN on Titan. *Bull. Amer. Astron. Soc.* **19**, 633.
- POYNTER, R. L., AND H. M. PICKETT 1982. *Submillimeter, Millimeter and Microwave Spectral Line Catalog*, JPL Publ. 80-23, rev. 1. Jet Propulsion Laboratory, Pasadena, CA.
- ROHART, F., D. DEROZIER, AND J. LEGRAND 1987. Foreign gas relaxation of the $J = 0 \rightarrow 1$ transition of HC^{15}N . A study of the temperature dependence by coherent transients. *J. Chem. Phys.* **87**, 5794–5803.
- YUNG, Y. L. 1987. An update of nitrile photochemistry on Titan. *Icarus* **72**, 468–472.
- YUNG, Y. L., M. ALLEN, AND J. P. PINTO 1984. Photochemistry of the atmosphere of Titan: Comparison between model and observations. *Astrophys. J. Suppl. Ser.* **55**, 465–506.

Article IV

The Structure, Stability, and Global Distribution of Io's Atmosphere

EMMANUEL LELLOUCH

DESPA, Observatoire de Paris, 92195 Meudon, France

MICHAEL BELTON

National Optical Astronomy Observatories, Tucson, Arizona 85726

IMKE DE PATER

University of California, Berkeley, California 94720

GABRIEL PAUBERT

IRAM, Granada, Spain

SAMUEL GULKIS

Jet Propulsion Laboratory, Pasadena, California 91109

AND

THÉRÈSE ENCENNAZ

DESPA, Observatoire de Paris, 92195 Meudon, France

Received February 27, 1992; revised May 6, 1992

Millimeter-wave observations of SO₂ have allowed the first groundbased direct detection of Io's neutral atmosphere. From observations of two SO₂ rotational lines, at 221.965 and 143.057 GHz, tentative detection of a third SO₂ line, at 346.652 GHz, and upper limits on two other lines, basic properties of Io's atmosphere are inferred. The SO₂ atmosphere appears to have a global temporal stability and can be represented by a collisionally thick 10¹¹–10¹² cm⁻³ atmosphere ($p = 3$ –40 nbar) covering a limited fraction (5–20%) of Io's surface, with possibly larger pressures on the trailing side than on the leading. The horizontal distribution of gaseous SO₂ is best described as the result of discrete distribution of (equilibrium or volcanic) sources rather than by vapor pressure equilibrium over a smooth distribution of surface frosts. The lower atmosphere seems surprisingly hot, about 500–600 K at 40 km. A reanalysis of the IRIS/Voyager observation of the ν_3 SO₂ band at 7 μ m over Loki, using a NLTE transfer model, suggests temperature/pressure conditions at Loki consistent with those derived for the global atmosphere from the millimeter-wave data. High temperatures in the lower scale height, however, are not accommodated by simple thermal models. Our results suggest that Io's atmosphere may be best described by a "volcanic source" atmospheric model, although some aspects of the "equilibrium" models,

notably the temporal stability, are also present. While the primary problem remains the need to unambiguously determine and explain the vertical thermal structure, it must be noted that if the atmosphere is hot, the concept of an atmosphere in dynamical equilibrium with one or more volcanic sources may provide a reasonable explanation to the Pioneer 10 occultation. Finally, new upper limits on atmospheric H₂S, SO, and CO were obtained. © 1992 Academic Press, Inc.

1. INTRODUCTION

Although the existence of a tenuous neutral atmosphere around Io was established almost two decades ago (Brown 1973), its fundamental properties—pressure, thermal structure, composition, spatial distribution, and temporal variability—still remain essentially unknown. Observations of the occultation of β Scorpii-C by Io (Bartholdi and Owen 1972, Smith and Smith 1972) provided an early, but model-dependent, upper limit of about 10⁻⁷ bar to the surface pressure near the terminator regions. The first (indirect) observational evidence for a relatively dense neutral atmosphere was obtained in 1973 with the detec-

tion of ionospheric layers, again over the terminator regions, by the Pioneer 10 radio-occultation experiment (Kliore *et al.* 1974, 1975). These measurements have turned out to be very difficult to use as a diagnostic tool for the lower atmosphere and there has arisen considerable controversy regarding the various interpretations (Johnson and Matson 1989). In 1979, near-IR spectra recorded by the Voyager IRIS experiment over one of the major volcanic hotspots discovered by Voyager, Loki, indicated the presence of gaseous SO₂, with a local pressure of about 10⁻⁷ bar, and allowed stringent upper limits to be placed on a number of other gases (Pearl *et al.* 1979). In the same year, solid sulfur dioxide (either as a frost or absorbate) was identified in groundbased near-IR spectra (Fanale *et al.* 1979, Smythe *et al.* 1979). These discoveries provided the first direct measurements of the pressure and composition of Io's atmosphere and set the stage for atmospheric models in which sublimation equilibrium, "buffering," plays a dominant role. Although the IRIS, groundbased, and Pioneer 10 observations related to different places and different epochs on Io, the various measurements have been used jointly to form the basis of aeronomic models (Kumar 1979, 1980, 1982, 1985, Kumar and Hunten 1982, Summers 1985). Kumar's models are particularly important for the work presented in this paper since they were the first to indicate that Io's atmosphere could be thermally inverted in the lowest scale height. It was this result that led us to the expectation that the millimeter-wave pure rotational spectrum of SO₂ would be in emission and ultimately led to the observations that we report below. Most post-Voyager attempts to observe Io's atmosphere have been carried out in the near-UV, notably with IUE (International Ultraviolet Explorer) (Bertaux and Belton 1979, Butterworth *et al.* 1980, Balster *et al.* 1990). Although none of these observations led to the actual detection of gaseous SO₂, it was possible to obtain upper limits on the global SO₂ amount that placed significant, if controversial (Belton, 1982), constraints on the amount of SO₂ in Io's atmosphere.

Of the several attempts to describe the horizontal distribution of gaseous SO₂, the models can be roughly classified as falling into "buffered," "volcanic," or "sputtered" categories. Within this categorization we have "thick (i.e., collisional)," "thin," "stable," and "transitory" varieties.

In the buffered models, the distribution of SO₂ reflects local vapor pressure equilibrium with the surface. Active sources and, in some models, atmospheric motions are ignored to the first approximation. These models have been mostly concerned with the location, state, and temperature of surface frosts that control the equilibrium pressure. Examples include the sublimation models considered by Pearl *et al.* (1979), Bertaux and Belton (1979), the regional (or albedo cold-trapping) buffered model of

Fanale *et al.* (1982), and the regolith (or subsurface) cold-trap model (Matson and Nash 1983). Since all of these models correspond to vapor pressure equilibrium at temperatures which vary by tens of degrees, the range of SO₂ pressures predicted by these models is enormous, from about 10⁻⁷ to 10⁻¹⁵ bar (Johnson and Matson 1989).

In "volcanic" models, the possible dynamic nature of Io's atmosphere and the role of volcanoes as sources are stressed. Kumar (1979) pointed out that the lifetime of an SO₂ atmosphere on Io must be very short, measured in days, due to the scavenging action of magnetospheric plasma flow around the satellite. He suggested that a continuous level of volcanic activity would be needed to maintain a stable atmosphere. Evidence for significant flows in the neutral atmosphere near the surface, related to major volcanic regions, has been documented by Lee and Thomas (1980). They find evidence in the Voyager images for flows such as those associated with Pele that extend to distances as great as 750 km. Buffered atmospheres must be inherently unstable due to large horizontal pressure gradients (Ingersoll *et al.* 1985), and such sublimation and volcanically driven flows (Ingersoll 1989, Baumgardner *et al.* 1987, Moreno *et al.* 1991) appear to be capable of distributing SO₂ across the surface to substantial distances. Whether or not the volcanic component of the atmospheric gas dominates in a particular locality depends on the strength of the source, the temperature and state of the frost at the surface, and the atmospheric temperature. Ultimately, however, the formal distinction between "equilibrium" and "volcanic" atmospheres must be somewhat artificial because equilibrium SO₂ frosts are losing mass that must be resupplied by the volcanoes (Kumar 1979, Ingersoll 1989).

In "sputtered" models, a tenuous atmosphere is produced by the impact of energetic magnetospheric particles with the surface (later generalized to include sputtering off an existing atmosphere; McGrath and Johnson 1987). This process produces a highly rarefied, collisionally thin, atmospheric corona that can feed new material into the Io torus (Lanzerotti and Brown 1983, Sieveka and Johnson 1985, 1986, Cheng *et al.* 1986). These processes may still be significant sources of neutral atmosphere in regions away from volcanic sources on the nightside where the surface temperature is well below 100 K.

In this paper, we show that the observations to date generally support the volcanic atmospheric concept, although aspects of all three types of models are probably at work. None of the current models, however, are found to satisfy the data completely. The primary problem appears to be in the calculation of the vertical structure of the atmosphere and, in particular, in the estimation of the atmospheric temperature in the first scale height. This was considered in detail by Kumar (1980, 1985) who found that near-UV heating was strong enough to produce a

temperature inversion down to the surface if the primary cooling was due to the fundamental vibrational bands of SO_2 . His models, like those of Johnson (1989), who considered only heating by plasma interaction, are quite hot. However, there is considerable uncertainty with these calculations because major cooling terms are omitted, e.g., cooling in the rotational bands and local cooling due to expansion of gas in the horizontal flow. Further uncertainties arise because the SO_2 infrared cooling was modeled only by analogy to the properties of CO_2 —a molecule that has different vibrational characteristics than SO_2 . Ingersoll *et al.* (1985) and Ingersoll (1989) avoided this problem by simply assuming an adiabatic lapse rate in the lowest scale height, while Moreno *et al.* (1991) included vibrational cooling but ignored UV heating. As a result, the models of these latter authors deal with extremely cold atmospheric flows that are limited in their ability to provide sizable atmospheric pressure at great horizontal distances from the sources of gas.

We report here on high-resolution millimeter-wave spectroscopic observations of Io performed in 1990 and 1991 that allowed the first unambiguous groundbased detection of SO_2 in the atmosphere of Io (Lellouch *et al.* 1990). The data are described and analyzed in terms of SO_2 atmospheric pressure, horizontal distribution, temperature, and temporal stability. We then present a reassessment of the IRIS–Voyager observations of the ν_3 band of SO_2 and show that non-LTE (local thermodynamic equilibrium) effects are important in the formation of the band. In this way, we are able to conclude that the global SO_2 atmosphere that we observed from Earth has characteristics (pressure, temperature) similar to those observed at Loki by Voyager. The paper concludes with a simple thermal model of Io's lower atmosphere and a discussion of new upper limits that we can place on the presence of H_2S , SO , and CO in Io's atmosphere.

2. OBSERVATIONS

2.1. General Procedure

Millimeter-wave observations of Io were conducted on the 30-m IRAM (Institut de Radio-Astronomie Millimétrique) radiotelescope located at Pico Veleta, Spain, and the 10.4-m CSO (Caltech Submillimeter Observatory) at Mauna Kea. We obtained data on five runs: UT 1990 January 16–18, 1991 January 12–14, and 1991 March 30–April 2 at IRAM, and UT 1991 January 1–3 and 23–25 at CSO. During each run, the maximum elongation of Io roughly coincided with the time of Jupiter transit. Moreover, because of the $42^{\text{h}}28^{\text{m}}$ rotation period of Io, observations on consecutive nights allowed us to study both the leading and the trailing sides of the satellite.

2.1.1. IRAM observations. Observations at IRAM

(Baars *et al.* 1987) can be performed using simultaneously three SIS (superconductor–insulator–superconductor) receivers, which operate respectively at 215–270 GHz (1.3 mm), 135–185 GHz (2mm), and 75–115 GHz (3mm) and can be used in either single or double sideband. Typical single sideband receiver temperatures are 220 K at 220 GHz, 140 K at 140 GHz, and 150 K at 115 GHz. The signal from each receiver can be analyzed in a backend (spectrometer) of 512 independent channels with 1-MHz resolution. For higher resolution studies, a 256×100 -kHz backend can also be used in parallel with the 1-MHz backend. The higher resolution analyzer can also be split into two equal parts (128×100 kHz) and shared between two receivers.

A facility for positional and velocity tracking of planetary satellites is not commonly found at radiotelescopes. At IRAM, one of us (G.P.) has implemented a position tracking program for satellites, using ephemerides from the Bureau des Longitudes, Paris, France (J.-E. Arlot and W. Thuillot, private communication). For velocity tracking, this program only corrects for the topocentric motion of the planet, and not for the planetocentric of the satellite. Since Io's orbital velocity changes rapidly, high resolution spectroscopic observations of sharp lines demand that Io's joviocentric velocity be corrected: Io's typical acceleration relative to Jupiter (projected on the Io–Earth line) is about 30 m/sec per minute (i.e., 24 kHz/min at 1.3 mm), which means that the line center of the 222-GHz line shifts by one FWHM (full width at half maximum, about 600 kHz; see below) in 25 min. Therefore, we structured all of our observations at IRAM as 2-min scans which, in the data reduction phase, were individually corrected for the joviocentric Doppler shift before they were coadded. This procedure, which leads to averaging scans with different central frequency and therefore different frequency sampling, slightly modifies the effective channel response function (see Appendix). The absolute accuracy of the ephemeris is of order ± 20 m/sec (15 kHz at 1.3 mm).

All our observations were done with the usual “position switch” mode, where the telescope alternatively points at the source and a reference sky position. At IRAM, a fast version of this mode was available (“wobbling”) which chopped the secondary of the telescope at a frequency of 1 Hz. The reference position was taken $240''$ ($10 R_J$) away from Io (the maximum allowed by the system), to minimize the contribution by Jupiter. Calibrations were performed with the standard hot load (chopper wheel)–cold load (sky) procedure. One calibration was made after a group of five scans at Io.

Pointing was checked every $\frac{1}{2}$ hr on Jupiter—and sometimes on Io itself—by azimuth and elevation drifts across the planet. Pointing corrections were generally of 1–2 arcsec, and never exceeded 4 arcsec. In addition, after 5

or 10 scans at Io, a scan on “anti-Io” (the sky position symmetric to Io with respect to Jupiter) was recorded, as a further check of the pointing and the calibration. Indeed, while the continuum of Io can in principle be detected and measured at millimeter wavelengths, the corresponding antenna temperature is small (a fraction of a K) because of a high dilution of Io ($1.1''$) in the main beam ($12''$ at 222 GHz), and can be severely offset (by several times its value) due to sky fluctuations, instrumental baseline ripples, and, when Io is several hours away from its maximal elongation position ($138''$), by the contribution of Jupiter in the sidelobes. Thus, the antenna continuum temperature does not provide a simple means to measure Io’s surface brightness temperature. In contrast, assuming the beam is symmetrical, the Io and anti-Io spectra are similarly affected, therefore the difference between the observed continua at Io and anti-Io is expected to be consistent with the known surface temperature of Io, providing another calibration and pointing control. With some exceptions, we merely rejected those scans which did not meet this continuum level criterion. The exceptions primarily occurred on the January 1991 run when Jupiter was near its node and the other Galilean satellites often fell into the beam with Io. This behavior was predictable and we were able to anticipate the observed changes in the continuum levels. In the case of this data we assumed that the entire emission line signal originated on Io. Frequency tuning was checked on interstellar sources (Orion, IRC + 10216 ...) and this provided extra control on the calibration.

2.1.2. CSO observations. CSO operates two receivers, one in the millimeter range (220–270 GHz) and one in the submillimeter range (330–360 GHz); only one of these receivers can be used at a time. Both receivers operate as a double sideband system, with noise temperatures of about 100 K for the millimeter and 200 K for the submillimeter bands. We only used the submillimeter receiver (Ellison *et al.* 1989). Spectral analysis is performed by means of two AOS (Acousto-Optical Spectrometers), covering 500- and 50-MHz intervals at spectral resolutions of 1 MHz and 100 kHz, respectively.

Positional tracking on Io was performed by using the general planetary software, where a polynomial fit is made through three planetary positions. For the planets, one usually takes the position once a day at 0^h UT. Since Io moves around Jupiter in about 42 hr, we made up an hourly ephemeris for the satellite to ensure proper tracking. There is no facility at CSO to track planets in velocity. We obtained the topocentric velocities of Io from M. Standish (JPL, private communication) at 10-min intervals. In contrast to the IRAM observations, we updated the observing frequency on line every 2 min, so that the “line” stayed at the center frequency. The JPL ephemeris is essentially the same as that from the Bureau des Longitudes.

Data were taken with the usual on-off method, where the off position was taken 125 arcsec ($5 R_J$) away from Io, in the direction opposite to Jupiter. Before each group of scans (but certainly before each change in observing frequency) a calibration scan was taken using a hot (chopper wheel) and cold (sky) load. The observed source temperature is corrected for atmospheric extinction and ground spillover. We measured the extinction by making skytips at the observing frequency. Typical values were 0.09 on 1990 Dec 31, 0.2–0.25 on 1991 Jan 1, 0.15 on Jan 2, and about 0.15–0.22 on 1991 Jan 23–25. In addition, the extinction was continuously monitored using a tipping radiometer at 225 GHz.

Pointing was checked roughly every half-hour on Jupiter, by using the five-point technique (points taken half a beam offset from the center position toward N, S, E, and W). The pointing was found to be generally within a few arcsec of the target. Frequency tuning was checked on Orion.

2.1.3. Calibration of antenna temperatures. The spectra shown in the figures are averages of individual scans that were summed after correction for orbital Doppler shifts. Since a calibration of Io’s continuum was not possible for the reasons mentioned above, a constant, or first-degree polynomial baseline has been subtracted from the data. The signal is obtained from the system in units of “antenna temperature” (T_A^*). This scale is fully corrected for the atmospheric extinction and can be converted to a brightness temperature (T_B) according to

$$T_A^* = [B_{\text{eff}}/F_{\text{eff}}]\eta T_B,$$

where $B_{\text{eff}}/F_{\text{eff}}$ is the ratio of the main beam to the forward efficiency and η is the “filling factor” of Io in the main beam. We assume the beam to be Gaussian, i.e., $\eta = 1 - 2^{-(D/\text{FWHM})^2}$ where D is Io’s angular size and FWHM is the full width at half maximum of the main beam. At IRAM, approximate values of FWHM are $13''$ at 222 GHz, $17''$ at 143 GHz, and $23''$ at 115 GHz, and the corresponding $B_{\text{eff}}/F_{\text{eff}}$ are about 0.50, 0.60, and 0.70. At CSO, the beam FWHM is $21''$ at 345 GHz and $B_{\text{eff}}/F_{\text{eff}}$ is 0.60.

2.2. Data

2.2.1. IRAM: UT 1990 January 16–18. Our first data set was obtained at IRAM on January 16 (leading side) and 17 (trailing side), 1990. We searched for SO₂ at 221.965 GHz (the (11,1,11)–(10,0,10) transition) and H₂S at 168.763 GHz (the (1,1,0)–(1,0,1) transition) at 1-MHz and 100-kHz resolution. SO₂ was unambiguously detected on both sides of Io, although the “continuum criterion” led us to reject approximately half of the scans and left us with admittedly low S/N data (Figs. 1a–1c). As mentioned

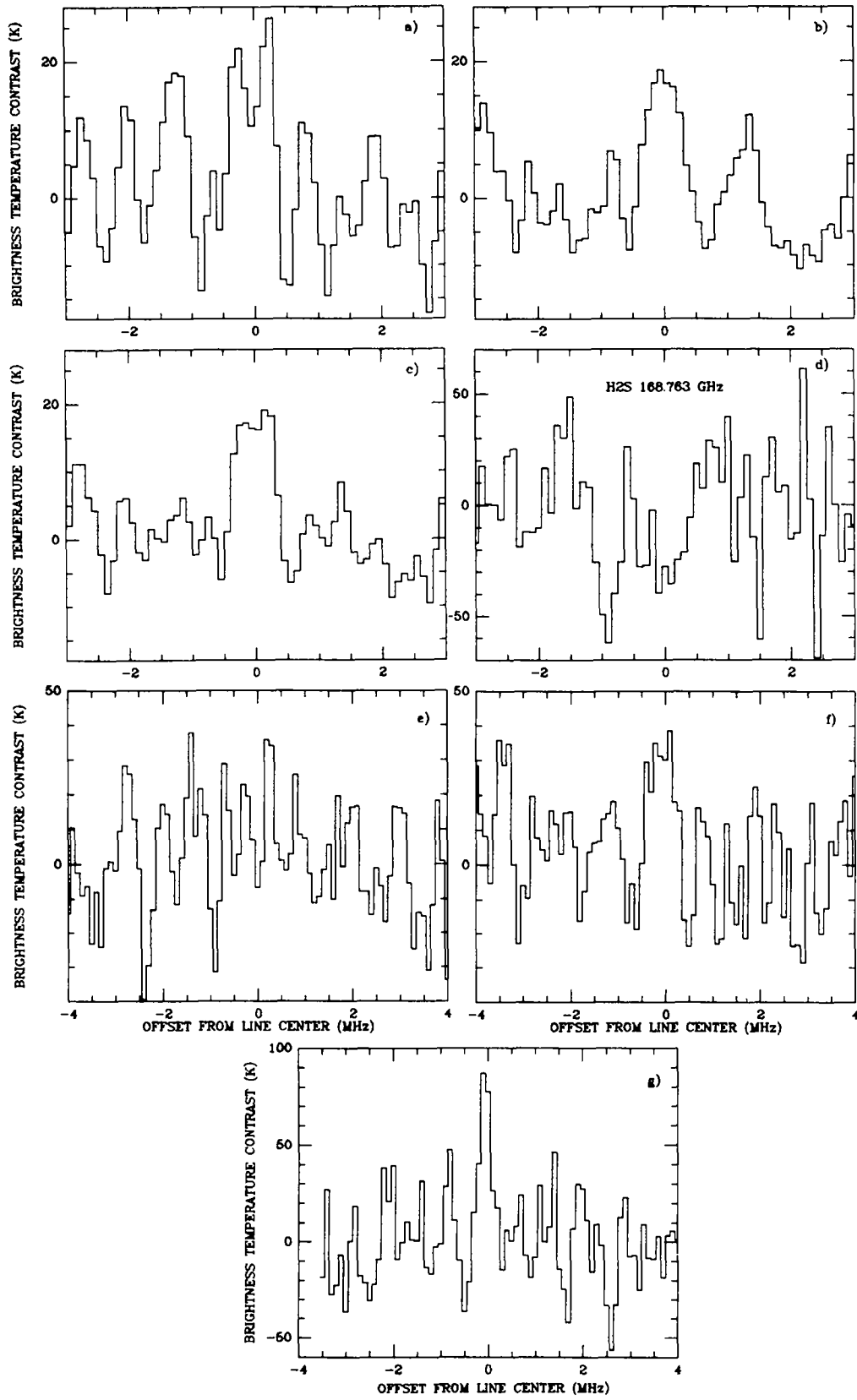


FIG. 1. IRAM observations UT 1990 January 16/18: (a–c) The SO₂ 221.965-GHz line. (a) Leading side, integration time $T_{\text{int}} = 79$ min. (b) Trailing side, $T_{\text{int}} = 129$ min. (c) Both sides, $T_{\text{int}} = 208$ min. (d) The spectrum near the 168.763-GHz line of H₂S (trailing side, $T_{\text{int}} = 129$ min). (e–g) Possible temporal variability of the SO₂ 221.965-GHz line. (e) Leading side UT 1990 January 17, 0.00–1.20 hr. $T_{\text{int}} = 32$ min. (f) Leading side UT 1990 January 17, 1.20–2.30 hr. $T_{\text{int}} = 35$ min. (g) First six scans on the trailing side UT 1990 January 17, 18.47–19.09 hr. $T_{\text{int}} = 12$ min.

in our preliminary report (Lellouch *et al.* 1990), the SO₂ line exhibits similar characteristics on both sides (a brightness temperature contrast (ΔT_B) of about 18 K and a FWHM of about 600 kHz. In addition (Fig. 1 of Lellouch *et al.* (1990) and Figs. 1e–1g hereafter), the SO₂ line seemed, at the 3σ confidence level, to show short term variability, which affected its contrast but not its width. (As the reader may note, the spectra shown in Fig. 1 are somewhat different from those published in Lellouch *et al.* (1990). This was due to a small error in the IRAM data reduction package regarding the averaging of scans with different frequency sampling.) In one instance, however, during the first 12 min of observation on the trailing side, we observed a much narrower (0.3 ± 0.2 MHz) and stronger (about 90 K) line. We speculated that this event may have been due to a possible cutoff of a volcanic source, shortly after our observations started on 1990 January 17 (UT about 19.00), which led to a rapid condensation of SO₂ onto the cold surface (Lellouch *et al.* 1990). Finally, the search for H₂S was negative, with a 1σ , 100-kHz upper limit of 29 K at Io (Fig. 1d).

2.2.2. CSO, UT 1990, December 30–1991 January 1. During the first run at CSO, we searched for SO₂ at 346.652 GHz [(19,1,19)–(18,1,18)]. Because the “scaling factor” ($B_{\text{eff}}/F_{\text{eff}} \times \eta$) is much smaller at CSO than at IRAM (1/800 vs 1/350), it is much more difficult to detect a line at CSO than at IRAM, although this line is intrinsically 3.5 times stronger than the 221.965-GHz line. The spectrum is shown in Fig. 2a (Hanning smoothed, unfolded spectrum). If the feature at the center is real, its reliability can be improved by using the folding technique SYMCOMET (Gerard 1987, Snyder *et al.* 1989) developed to detect weak cometary lines. One folds the spectrum with respect to a center channel, which thus increases the S/N in the spectrum, except at the center channel. Since the line is expected to be at channel 256, we folded around this channel; the result is shown in Fig. 2b. The rms noise in this spectrum is 20 mK, or 16 K on Io. We inspected similar spectra, folded over different channels. We typically detected a line, although with a poor S/N ratio, if the center channel was chosen to lie between channels 250 and 259, and did not detect a line when a channel outside this range was chosen as the center channel. From these numerical experiments, we conclude that we may have detected the line at the 2σ level: the strength of the feature in Fig. 2b is 40 mK, or 32 K on Io, a factor of 2 above the rms background noise. The width of the line is about 1 MHz.

2.2.3. IRAM: UT 1991 January 12–14. The main goal of this run was to repeat the January 1990 observation to study the temporal variability of Io’s SO₂ atmosphere, on both long time scales (1 year) and short term periods (hours). This run occurred during the period of the satel-

lites’ mutual events, which caused the continuum on Io or anti-Io to be occasionally contaminated by contribution of one of the other satellites in the beam. Because this happened in a predictable way, however, the “continuum criterion” could still be used. Thanks to excellent sky conditions, the 221.965 line was easily detected at 100-kHz resolution, on both sides, with an improved S/N (Figs. 3a–3c). On average, the line features (contrast and width) remained unchanged compared to the previous year. In addition, on January 13/14, the 2-mm receiver, which was unavailable on January 12/13, could be operated and allowed us to detect the (16,2,14)–(16,1,15) line at 143.057 GHz on the trailing side (Fig. 3d), in spite of a worse “scaling factor” (1/474 vs 1/371 at 222 GHz). The 143-GHz line appeared somewhat narrower than the 222-GHz line, as expected for pure Doppler-broadened lines (see below). The 222-GHz line appeared slightly broader on the trailing than on the leading side (Fig. 3). Interestingly, both lines were also detected in the 1-MHz backend, although they were unresolved. As opposed to the 1990 data, the lines detected on 1991 January 12–14 did not show any obvious temporal variability, even at a low confidence level. The 222-GHz line on the trailing side seems to show some asymmetry, with one wing steeper than the other; however, this is within the noise level and we cannot think of any obvious mechanism that would lead to an asymmetry in the line profile. Finally, using the 3-mm receiver in parallel we searched for CO at 115 GHz, but only at 1-MHz resolution. This line was not detected, with a 1σ upper limit of 3 K at Io at 1-MHz resolution.

2.2.4. CSO: UT 1991 January 23–25. During the second run at CSO, we searched for SO at 346.529 GHz. The line was not detected, despite our efforts at smoothing and folding. We obtained an rms noise of 23 K at Io at 100-kHz resolution or 16 K after folding the spectrum (Figs. 2c and 2d).

2.2.5. IRAM: UT 1991 March 31–April 2. The last run at IRAM was devoted to a search for weaker SO₂ lines, in an attempt to better constrain Io’s thermal structure. The idea is that the 222-GHz line is saturated to some extent, depending on the temperature (Lellouch *et al.* 1990). The lower the temperature, the more saturated the line must be. Detection of much weaker lines would indicate a strong saturation of the 222-GHz line, and therefore a low atmospheric temperature, whereas nondetection of reasonably weaker lines would be evidence for high atmospheric temperatures. Our observations, unfortunately, were seriously hampered by tuning and stability problems on the 1.3-mm receiver. During the first night (leading side), we searched for the 146.550-GHz line ((10,4,6)–(11,3,9)) at 100-kHz resolution, which (at 300 K) is intrinsically 8 times weaker than the 143-GHz line. The 1-MHz backend was centered so as to observe this line

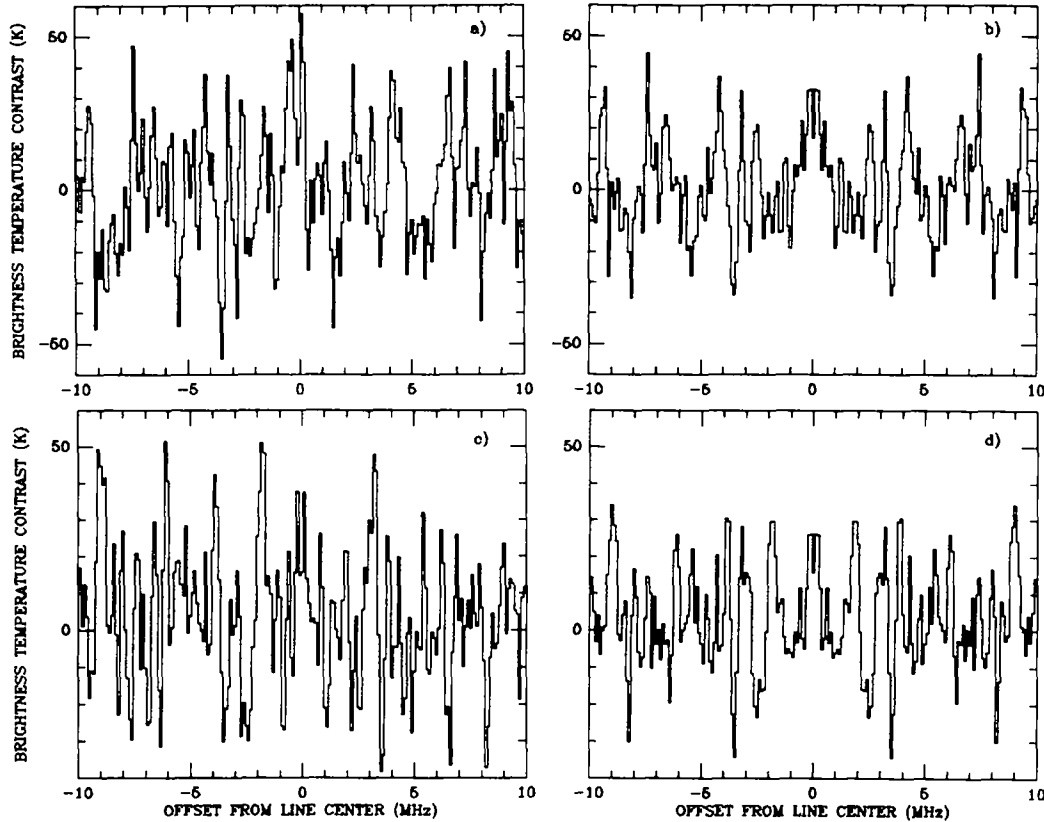


FIG. 2. CSO observations. (a) The SO_2 346.65-GHz line, $T_{\text{int}} \sim 16$ hr (UT 1990 December 31–1991 January 1). (b) Same spectrum, after folding. (c) Spectrum near the 346.53-GHz line of SO, $T_{\text{int}} \sim 8$ hr (UT 1991 January 23–24). (d) Same spectrum, after folding.

simultaneously with the (4,2,2)–(4,1,3) 146.605-GHz line (3 times weaker than the 143-GHz line). No line was detected (Fig. 4a) and on the following night we centered the 100-kHz backend on the 146.605-GHz line, which was not detected either (Fig. 4b). Since these nondetections appear of crucial importance for the interpretation in terms of atmospheric temperatures (see below), we have very carefully checked possible instrumental problems. Observations of interstellar sources showed that the 2-mm receiver was tuned to the right frequency, remained stable, and did not experience any particular calibration problems. The pointing and tracking at Io was good, as attested by the continuum level test (Fig. 4c). We finally strongly believe that on March 31/April 1, 1991, the SO_2 146.550 and 146.605 lines did not show up at the 7.9 and 5.7 K (1σ) level.

3. ANALYSIS

3.1. Local Thermodynamical Equilibrium

The Einstein coefficient A of the millimeter-wave transitions can be computed as a function of their absolute strength S (in cm/mole) according to

$$A = \frac{8\pi k T \nu}{h} \frac{Q}{g} e^{hcE/kT} S,$$

where T is temperature, Q the partition function at temperature T (5898.9 at 300 K), and ν , E , and g are the transition frequency (cm^{-1}), lower energy level (cm^{-1}), and degeneracy, respectively. The emission is in local thermodynamical equilibrium (LTE) as long as this transition probability is small compared with the collisional deactivation probability $k_d \times N$, where k_d is the relaxation deactivation rate ($\text{cm}^3 \text{sec}^{-1}$) and N the atmospheric concentration. Bass *et al.* (1971) have measured a rotational relaxation time of $3 \times 10^{-10} \text{ sec/atm}$ at 300 K, which gives $k_d = 1.4 \times 10^{-10} \text{ cm}^3 \text{sec}^{-1}$. Using spectroscopic information from the JPL catalog (Poynter and Pickett 1982), Table I computes, for the various transitions we have observed, the Einstein coefficient and the critical concentration $N_0 = A/k_d$. This shows that the rotational transitions of SO_2 are in LTE in most of Io's neutral atmosphere: for a surface concentration of about $2 \times 10^{11} \text{ cm}^{-3}$, which, as we show below, is inferred from our measurements, and a typical critical concentration of $2 \times 10^5 \text{ cm}^{-3}$, the lines are in LTE up to about 14 scale heights. In fact, since the

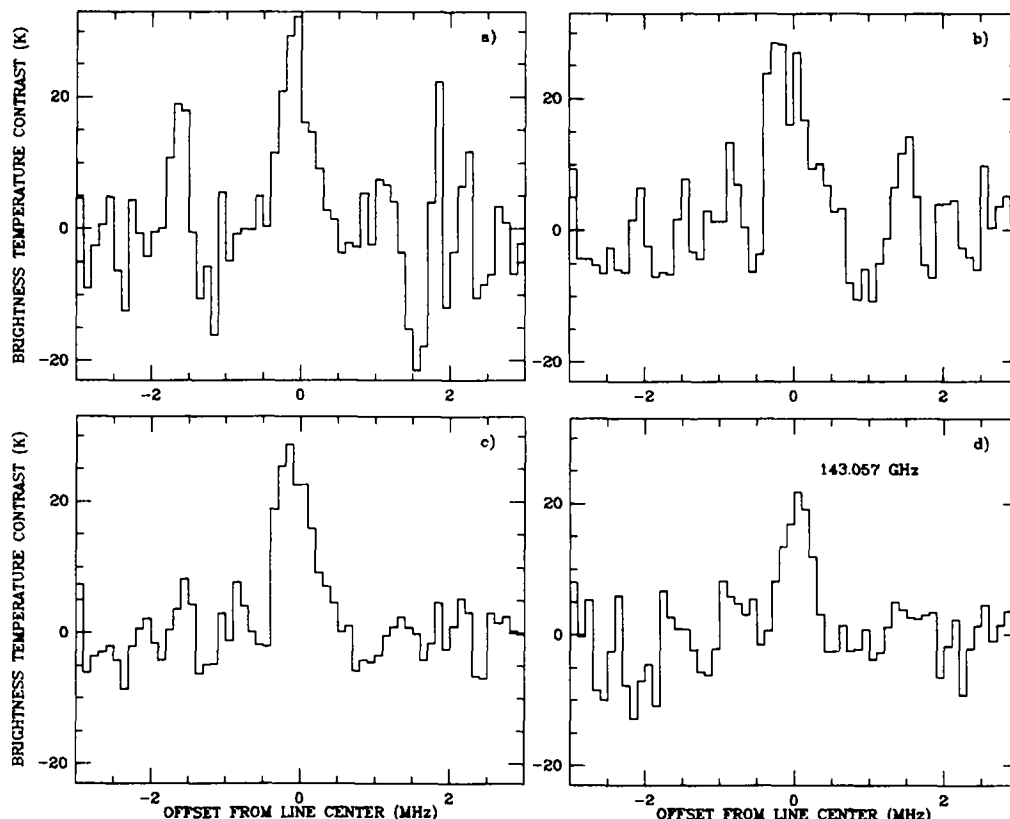


FIG. 3. IRAM observations, UT 1991 January 12–14. (a–c) The SO_2 221.965-GHz line. (a) Leading side, $T_{\text{int}} = 126$ min. (b) Trailing side, $T_{\text{int}} = 193$ min. (c) Both sides, $T_{\text{int}} = 319$ min. (d) The SO_2 143.057-GHz line on the trailing side, $T_{\text{int}} = 158$ min.

line centers are optically thin at the critical level (τ of order 4×10^{-5} for the 222-GHz line), the departure from LTE does not impact at all on the computation of the emission.

3.2. Broadening Regime

Smith and Smith (1972) obtained an upper limit for the pressure of Io's atmosphere from the β Scorpil occultation in 1971. Due to different molecular refractivities, this upper limit depends on the nature of the dominant atmospheric species, and is, for example, equal to $0.09 \mu\text{bar}$ for N_2 , $0.13 \mu\text{bar}$ for CH_4 , and $44 \mu\text{bar}$ for H_2 , assuming an isothermal atmosphere at 100 K. Since the pressure limit scales as $T^{2.5}$, more conservative upper limits, assuming $T = 500$ K, are obtained by multiplying these numbers by 50. As is characteristic for a stellar occultation, these numbers apply to Io's limb. Since the above species do not condense out at Io's surface temperatures, the corresponding limits should hold at any location on Io. If the dominant gas in Io's atmosphere is condensible, the upper limits above cannot be used. However, more stringent upper limits (of order 10^{-8} bar) have been

inferred for a number of common gases (CO_2 , CH_4 , NH_3 , H_2O , H_2S ...) from IRIS/Voyager data at the hot spot (Loki) where SO_2 was detected. Therefore, excluding the hydrogen atmosphere case as implausible based upon Jeans escape criteria, a maximum ground pressure of about $5 \mu\text{bar}$ must apply at any location on Io (except possibly in localized volcanic plumes).

The Doppler FWHM at 222 GHz is approximately 220 kHz at 120 K, whereas the Lorentz collisional width is of order $6 \text{ kHz}/\mu\text{bar}$. Thus, the SO_2 transitions are purely Doppler broadened.

3.3. First-Order Estimate of the SO_2 Pressure Distribution

Unless the characteristic atmospheric temperature is about 900 K, the observed FWHM are larger than the Doppler width, indicating saturation. This provides a straightforward, analytic way to infer at first order the SO_2 amount on Io. Considering a surface with temperature T_s (130 K) and microwave emissivity ϵ (0.9), a fraction ρ of which is covered by an isothermal atmosphere at temperature T_{atm} , the line contrast at frequency ν from line center is

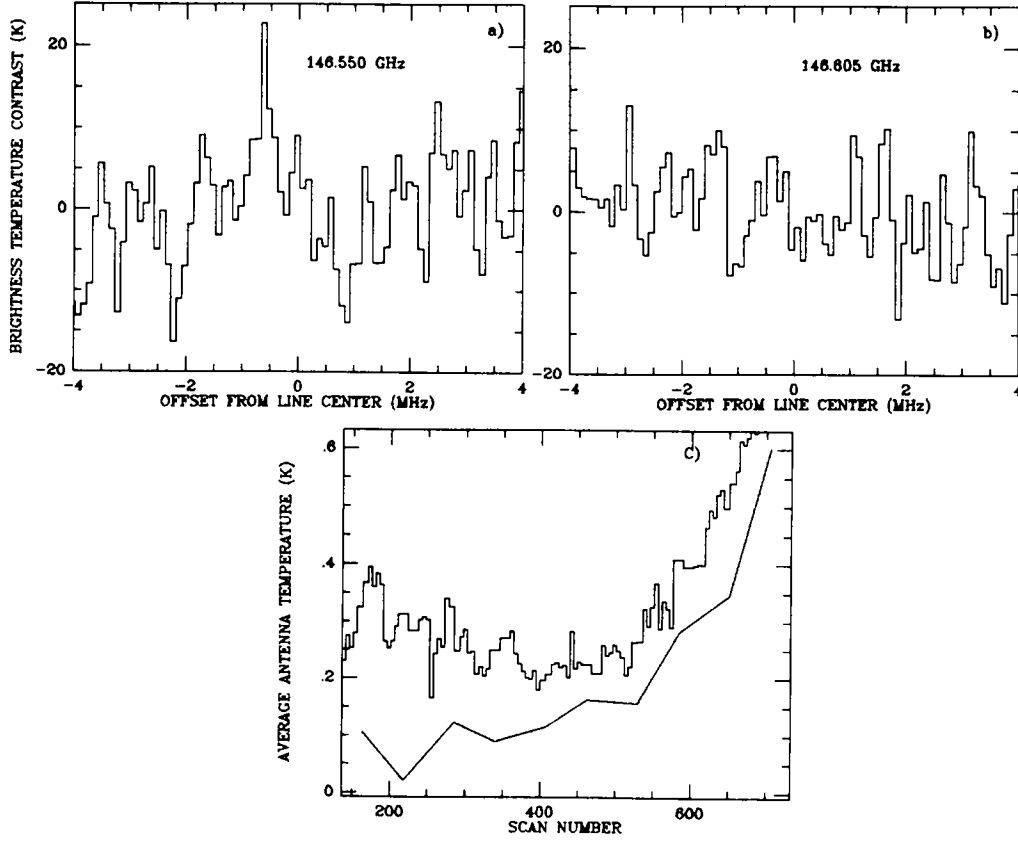


FIG. 4. IRAM observations, UT 1991 March 31–April 2. (a) The spectrum near the 146.550-GHz line of SO₂. $T_{\text{int}} = 6$ hr. (b) The spectrum near the 146.605-GHz line of SO₂. $T_{\text{int}} = 4$ hr. (c) The average (continuum) temperature on Io (histograms) and “anti-Io” (solid line) during the 146.605-GHz observations, expressed in antenna temperature. Io’s theoretical continuum is about 0.22 K in this scale. The increase of continuum level with scan number is due to the progressive filling of Jupiter in the sidelobes.

$$\Delta T_B(\nu) = \rho(T_{\text{atm}} - \epsilon T_S)(1 - e^{-\tau(\nu)}) \quad (1)$$

and the half width at half maximum (HWHM) $\Delta\nu$ is defined by

$$\Delta T_B(\Delta\nu) = 1/2(\Delta T_B(0)) \quad (2)$$

or

$$1 - e^{-\tau(0)} = 2(1 - e^{-\tau(\Delta\nu)}), \quad (3)$$

where $\tau(\nu)$ is the integrated vertical opacity at frequency ν . For a Gaussian profile,

$$\tau(\nu) = \tau(0) \times 2^{-(\nu/\alpha_{\text{Dop}})^2} \quad (4)$$

where α_{Dop} is the Doppler HWHM.

From the measurement of $\Delta\nu$, Eqs. (3) and (4) allow us to estimate $\tau(0)$ and to infer the SO₂ column density \mathcal{N} (mole cm⁻²), since

$$\tau(0) = \frac{S(T_{\text{atm}})\mathcal{N}}{\sqrt{\pi}\alpha_{\text{Dop}}},$$

where $S(T_{\text{atm}})$ is the absolute line strength at temperature T_{atm} . \mathcal{N} can be expressed in terms of a partial pressure p (bar), assuming an airmass of 1, from $\mathcal{N} = pHL$, where H is the atmospheric scale height (cm) and $L = 2.687 \times 10^{19}$ cm⁻². Assuming an airmass of 1 will be justified by our finding that the atmosphere is, at least partly, located in the subsolar regions. Finally, once $\tau(0)$ is known, ρ is derived from (1).

To produce a line with contrast ΔT_B , the atmospheric temperature must be at least equal to $T_{\text{atm}} = \epsilon T_S + \Delta T_B$. This gives a lower limit $T_{\text{atm}} \geq 140$ K for $\Delta T_B = 23$ K. τ , p , and ρ strongly depend on T_{atm} through α_{Dop} and Eq. (1), and typical values are given in Table II.

Although this analysis is really rudimentary, it provides a hint of the basic features of Io’s SO₂ atmosphere. The atmospheric pressure is from one to a few tens of nanobars, and since ρ is significantly less than 1 except in the extreme temperature cases (very cold or very hot), this atmosphere probably covers only some 10% of the satel-

TABLE I
Local Thermodynamical Equilibrium of Several SO₂ Rotational Lines

Frequency (GHz)	Absolute strength (cm/mole)	Degeneracy	Energy (cm ⁻¹)	Einstein coefficient (sec ⁻¹)	Critical N_0 (cm ⁻³)
346.652	1.11×10^{-21}	39	105.30	5.05×10^{-4}	3.6×10^6
221.965	3.18×10^{-22}	23	34.55	1.12×10^{-4}	8.2×10^5
143.057	1.73×10^{-22}	33	90.82	3.58×10^{-5}	2.6×10^5
146.605	4.66×10^{-23}	9	8.33	2.44×10^{-5}	1.8×10^5
146.550	2.05×10^{-23}	21	57.55	5.82×10^{-6}	4.3×10^4

lite's hemisphere facing Earth. This led us to support the idea proposed earlier (e.g., Bertaux and Belton 1979) that the SO₂ atmosphere of Io is in equilibrium with the surface temperature, leading to a significant atmosphere in the subsolar regions only (Lellouch *et al.* 1990). In addition, since pressures we derive in the vicinity of the subsolar region are lower than the equilibrium values for the average noon temperature (130 K, i.e., 1.3×10^{-7} bar, assuming Wagman's (1979) analysis of SO₂ vapor equilibrium data), this qualitatively favors the "regional" buffered models proposed by Fanale *et al.* (1982), where the SO₂ abundance is controlled by the small scale brightest and coldest areas. However, as shown below, further analysis of our data shows that the reality is probably more complex than this simple picture.

3.4. Isothermal Fits

The simple analysis above, which only makes use of the observed line contrast and FWHM, provides the SO₂ amount and surface coverage as a function of the atmospheric temperature, but does not give any information on the temperature, except for the loose constraint $140 \text{ K} \leq T_{\text{atm}} \leq 900 \text{ K}$. In fact, it is possible to derive some constraints on the temperature, by fitting the observed lineshapes of the detected 222-, 143-, and 346-GHz lines and by using the upper limits on the 146.550- and 146.605-

GHz lines. By "temperature," we mean here the average temperature of the lower atmosphere (about 0–50 km), from which the bulk of the observed emission comes (see below).

In the framework of isothermal atmosphere models, we have generated spectra of all the observed (detected or not) lines, for a variety of atmospheric temperatures ($120 \text{ K} \leq T_{\text{atm}} \leq 2000 \text{ K}$) and SO₂ number densities at the surface ($5 \times 10^9 \text{ cm}^{-3} \leq N \leq 10^{13} \text{ cm}^{-3}$). For each line, 13 (temperature) \times 18 (number density) spectra were computed. Comparison of models with observations was performed using the χ^2 test, which defined a best model fit (minimum of χ^2) and an acceptable model domain at the $n\sigma$ confidence level (defined by $\chi^2 \leq \chi^2_{\text{min}} + n^2$). The atmospheric fractional coverage was kept as a free parameter, but was constrained to be constant when fitting a set of different spectra.

3.4.1. Year-to-year temporal variability. We first fitted independently the January 1990 data (the 222-GHz line on both sides—two spectra) and the January 1991 data (the 222-GHz line on both sides plus the 143-GHz line on the trailing side—three spectra). Figure 5a shows the acceptable model spaces. For 1990, the best fits are obtained for the lower temperatures ($135 \text{ K} \leq T_{\text{atm}} \leq 300 \text{ K}$ at 1σ), as qualitatively indicated by the "saturated" shape of the 222-GHz line in Fig. 1. A typical solution is $T_{\text{atm}} = 200 \text{ K}$, $N = 5 \times 10^{11} \text{ cm}^{-3}$ (or $p = 14 \text{ nbar}$). For

TABLE II
First-Order Analysis of Millimeter-Wave Lines

Line/obs. date	FWHM (kHz)	Contrast (K)	T_{atm} (K)	Line center opacity	Pressure (nbar)	Fractional coverage
222 GHz (ave.)	600	20	140	63	5.3	0.87
			300	5.7	4.1	0.11
			500	2.0	7.8	0.06
			900	.04	1.3	0.74
222 GHz (1/13/91)	540	24	300	3.7	2.7	0.13
222 GHz (1/14/91)	780	21	300	24	18	0.11
143 GHz (1/14/91)	520	19	300	31	27	0.10
346 GHz (ave.)	1000	32	300	7.6	2.5	0.18

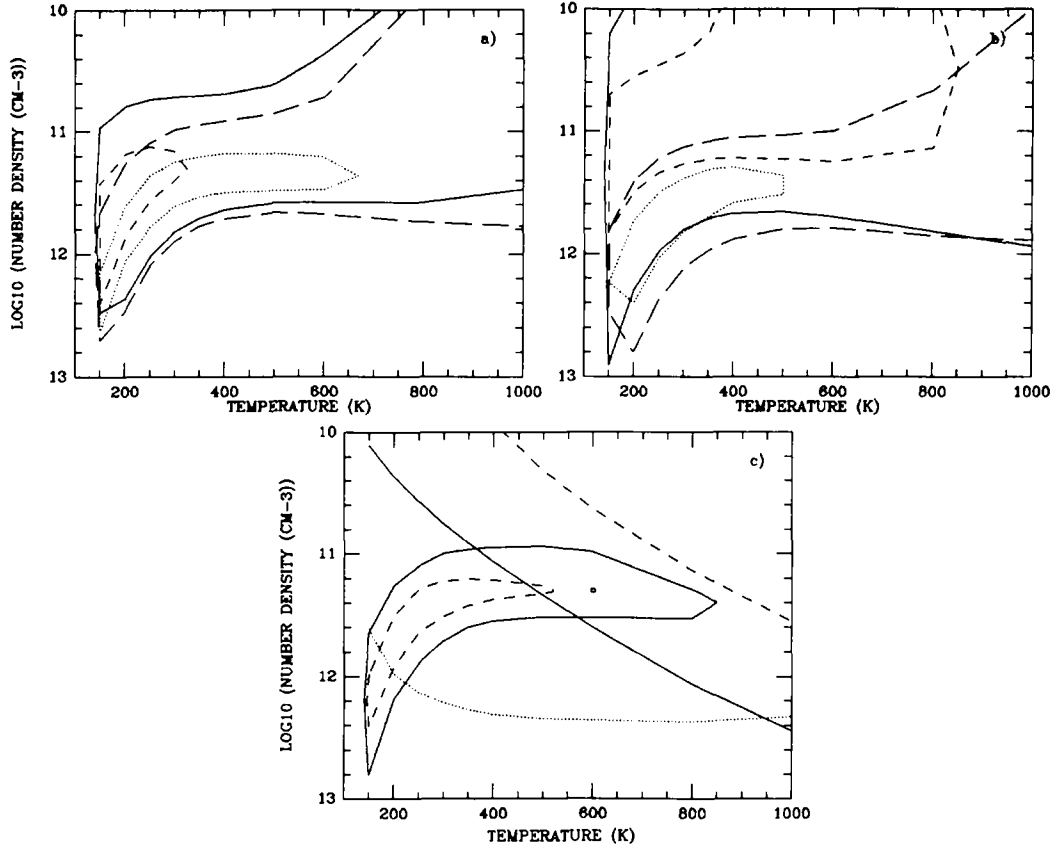


FIG. 5. Acceptable model spaces in the (temperature, number density) plane for fitting of various groups of data. (a) Independent fitting of the 1990 and 1991 SO_2 spectra. Short dashed, 1990, 1σ confidence. Solid, 1990, 2σ . Dots, 1991, 1σ . Long-dashed, 1991, 2σ . (b) Independent fitting of the leading and trailing side data of January 1991. Short dashed, leading side, 1σ confidence. Solid, leading side, 2σ . Dots, trailing side, 1σ . Long-dashed, trailing, 2σ . (c) Simultaneous fitting of SO_2 data. Circled lines, 1σ (dashed) and 2σ (solid) lines defined by the fit of the detected 222-, 143-, and 346-GHz lines. Oblique lines, acceptable 1σ (dashed) and 2σ (solid) defined by the nondetection of the 146.55- and 146.605-GHz lines. The acceptable domain lies above these lines. Dotted line, line defining the disk-averaged SO_2 column density equal to the upper limit of Ballester *et al.* (1990), rescaled to an airmass of 2 (0.015 cm-am, see text). The domain below this line violates this upper limit.

1991, the best fit is obtained at $T_{\text{atm}} = 400$ K, $N = 2 \times 10^{11} \text{ cm}^{-3}$ ($p = 11$ nb), but the 1σ domain extends from 150 to 600 K. The 1σ 1990 and 1991 spaces significantly overlap and correspond to indistinguishable atmospheric fractional coverages, as qualitatively expressed by the consistent line contrast of the 222-GHz line observed in 1990 and 1991. We conclude that the atmosphere remained globally unchanged between January 1990 and January 1991, and take this as a strong indication that Io's atmosphere has a stable, time-independent component.

3.4.2. Leading vs trailing sides. We next compare the data obtained on the leading and trailing sides observed respectively on 1991 January 13 (the 222-GHz line) and 14 (the 222- and 143-GHz lines). Figure 5b shows the corresponding model spaces. In both cases, the atmospheric temperature range extends from 150 to about 600 K. However, while the best fit for the leading side is

obtained for $T_{\text{atm}} = 350$ K, $N = 7 \times 10^{10} \text{ cm}^{-3}$ (i.e., $p = 3$ nbar), the density providing the best agreement on the trailing side is $N = 3 \times 10^{11} \text{ cm}^{-3}$ for $T_{\text{atm}} = 350$ K (i.e., $p = 14$ nbar). Model spaces for the leading and trailing sides do not intersect at the 1σ level, and the case for a larger average density on the trailing side can be made at approximately the 70% confidence level. In contrast, the fractional atmospheric coverage is about the same on both sides (e.g., at $T = 350$ K, $p = 11\%$ on the leading and 9% on the trailing side).

3.4.3. Atmospheric temperature. To constrain the atmospheric temperature, we use our entire data set and assume that it describes a temporally stable atmosphere, identical on the leading and trailing sides. We first fitted simultaneously all the detected lines (the 222-GHz line observed four times, plus the 143-GHz and the 346-GHz lines) by a single model—i.e., a single atmospheric tem-

perature, number density, and coverage. The best fit is obtained for $T_{\text{atm}} = 300$ K, $N = 2 \times 10^{11} \text{ cm}^{-3}$ (or $p = 9$ nbar), $\rho = 0.12$, average column density = 0.002 cm-am. Although low temperatures provide the best fit, the 1- and 2 σ temperature domains correspond to $135 \text{ K} \leq T_{\text{atm}} \leq 500$ K and $135 \text{ K} \leq T_{\text{atm}} \leq 800$ K, respectively (Fig. 5c). The upper limits on the 146.550 and 146.605 lines, for their part, provide a lower limit on the temperature. Synthetic spectra of the 146-GHz lines were generated on the same (T_{atm}, N) grid, using the atmospheric coverage found in each (T_{atm}, N) case from the best fit of the 222-, 143-, and 346-GHz lines. We then computed the 1- and 2 σ (T_{atm}, N) domains from the point of view of the upper limits on the 146-GHz lines. As is evident from Fig. 5c, at 1 σ , this domain and the one obtained previously from the shape of the detected lines are mutually exclusive. A compromise can be found at the 2 σ level, corresponding to the ranges $500 \text{ K} \leq T_{\text{atm}} \leq 700$ K and $1 \times 10^{11} \leq N \leq 3 \times 10^{11} \text{ cm}^{-3}$. The best compromise is obtained for $T_{\text{atm}} = 600$ K, $N = 2 \times 10^{11} \text{ cm}^{-3}$ ($p = 16$ nbar), $\rho = 0.05$, and an SO_2 column density of 0.0017 cm-am averaged over the disk. With this model, the 222-GHz line has a line center opacity of about 2.4 at the surface, and the $\tau = 1$ level is reached at about 40 km altitude, so that the “characteristic” atmosphere temperature we use is actually relevant to this altitude region.

Figure 6 shows synthetic spectra superimposed on the observations. The dotted lines show the spectra resulting from this final “baseline” model, whereas the dashed lines specifically show the best model fits for each individual observation. It is worth repeating that the final “baseline” model is obtained by fitting all the data against a single atmosphere model and that, in particular, it assumes that the stringent upper limits on the 146-GHz lines are real and not due to experimental problems that are not understood nor to intrinsic and rapid atmospheric variability.

3.5. “Smooth” Vapor Pressure Equilibrium Models

Most vapor pressure equilibrium models ignore the exact distribution of SO_2 frost and merely assume a horizontal distribution of equilibrium temperature T_{eq} , which smoothly varies between the subsolar and the terminator regions. The partial pressure of SO_2 in Io’s atmosphere is therefore described as a simple analytic function of a

limited number of variables (the solar zenith angle (SZA) and sometimes the latitude). Examples of such models and associated functions are given in Ballester *et al.* (1990). The dependence of the temperature on SZA varies from one model to another, reflecting the uncertainty on the dominant mechanism controlling the frost temperature. We attempted to see whether this kind of “smooth” buffered model is permitted by the data. Rather than testing the different models proposed in the literature, we focused on a certain function of $T_{\text{eq}}(\text{SZA})$. We assume that the surface temperature of the frost is defined by instantaneous equilibrium with the sum of the local solar insolation and some kind of “internal” heat flux ϕ_{int} , which defines the nightside temperature T_N . This results in

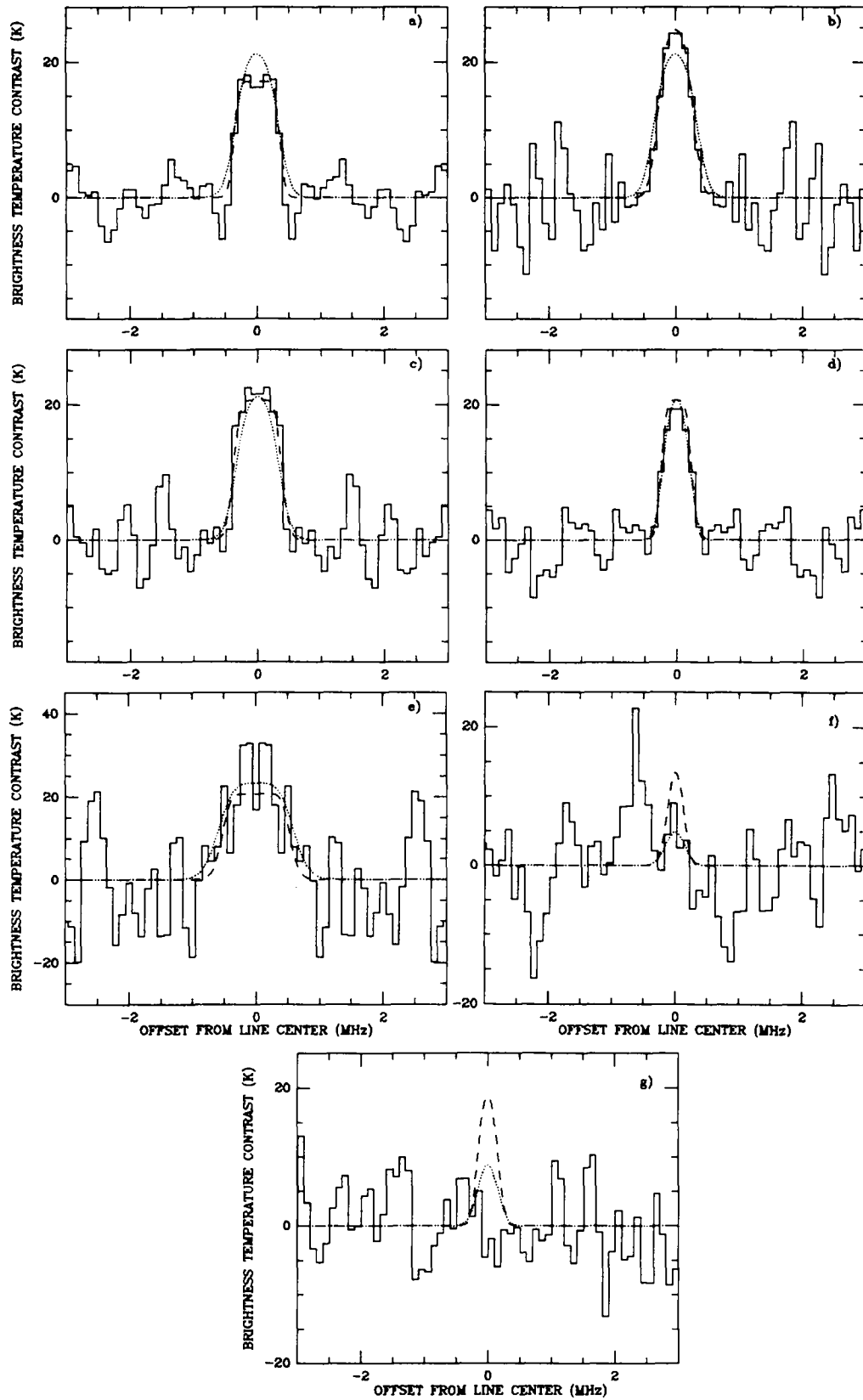
$$\sigma T(\theta)^4 = \phi_{\text{int}} + \phi_{\text{sol}} \cos(\theta),$$

where ϕ_{sol} is the incoming solar flux at zenith, θ is the solar zenith angle, and $\phi_{\text{int}} = \sigma T_N^4$. This can be expressed as

$$T(\theta) = T_N \left(1 + \left[\left(\frac{T(0)}{T_N} \right)^4 - 1 \right] \cos(\theta) \right)^{1/4}.$$

The SO_2 partial pressure is then a function (via Wagon’s (1979) equation) of SZA, with the temperature of the subsolar point and the night side hemisphere as free parameters. The main problem for our purpose is that the characteristic temperature of the lower atmosphere, or more specifically the vertical thermal structure, must also be specified as a function of SZA. Simple thermal models (see below) show that if the atmospheric heating is dominated by solar radiation, then the combination of a lower surface pressure and the heating dependence on SZA makes the atmosphere much cooler toward the terminator than at the subsolar point. More precisely, in this case, assuming the Fanale *et al.* (1982) surface temperature model, the atmospheric temperature remains at any altitude within 10 K of the surface temperature at $\text{SZA} \geq 60^\circ$ (where $T_{\text{eq}} \leq 108$ K, i.e., $p \leq 10^{-10}$ bar). In contrast, if the heating is mainly due to magnetospheric precipitation, then, for reasonable interaction across sections

FIG. 6. Fits of the 222-, 143-, and 346-GHz SO_2 lines for various models. The data have been folded and are shown as histograms. The dots show the “final best-fit” model (see text): $T_{\text{atm}} = 600$ K, number density $N = 2 \times 10^{11} \text{ cm}^{-3}$ ($p = 16$ nbar), atmospheric coverage $\rho = 5\%$. The dashed lines show other models specifically adapted to the individual observations. (a) Average 222-GHz line observed in 1990. Dashed line: $T_{\text{atm}} = 200$ K, $N = 5 \times 10^{11} \text{ cm}^{-3}$ ($p = 14$ nbar), $\rho = 22\%$. (b) Same line observed on 1991 January 13 (leading side). Dashed: $T_{\text{atm}} = 350$ K, $N = 7 \times 10^{10} \text{ cm}^{-3}$ ($p = 3$ nbar), $\rho = 11\%$. (c) Same line on January 14 (trailing side). Dashed: $T_{\text{atm}} = 350$ K, $N = 3 \times 10^{11} \text{ cm}^{-3}$ ($p = 14$ nbar), $\rho = 9\%$. (d) Same as (c) for the 143-GHz line. (e) 346-GHz observed line. Dashed line corresponds to the best-fit model of the detected lines only (see text): $T_{\text{atm}} = 300$ K, $N = 2 \times 10^{11} \text{ cm}^{-3}$ ($p = 9$ nbar), $\rho = 12\%$. (f) Same as (e) for the undetected 146.550-GHz line. (g) Same as (e) for the undetected 146.605-GHz line.



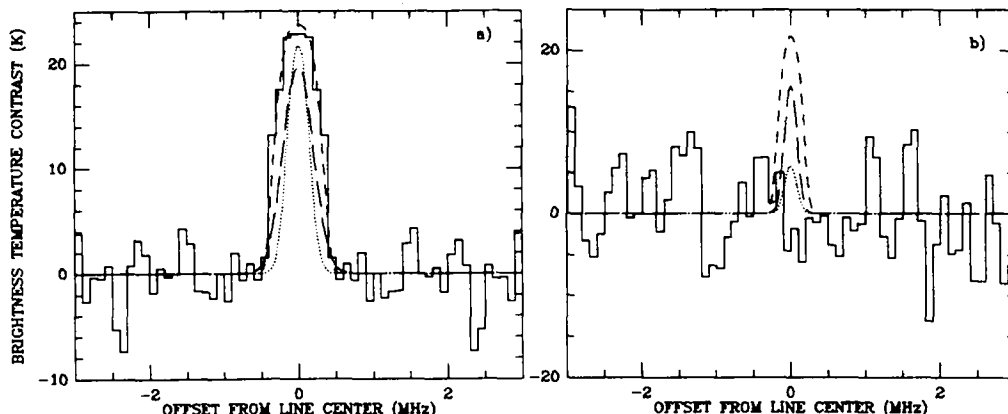


FIG. 7. Attempts to fit the 222-GHz detected (a) and the 146.605-GHz undetected (b) lines with different "smooth" vapor equilibrium models (see text). Histograms: data. Short-dashed line: $T_{\text{atm}} = 136$ K, $T(0) = 135$ K, $T_N = 100$ K. Dotted line: $T_{\text{atm}} = 200$ K, $T(0) = 110$ K, $T_N = 90$ K. Long-dashed line: $T_{\text{atm}} = 500$ K, $T(0) = 113$ K, $T_N = 93$ K.

(10^{-16} cm 2), the lower atmosphere is more efficiently warmed for SO $_2$ column densities of about 10^{16} cm 2 , i.e., away from the subsolar regions. If the main source of energy is located on Io (e.g., volcanoes), then the atmospheric temperature is not expected to be correlated with SZA. Because the reality could be a mixture of all these situations, and because all horizontal temperature contrasts may be washed out by winds (Ingersoll *et al.* 1985), we adopted a constant atmospheric temperature T_{atm} , independent of SZA, although we admit this could be very far from the truth.

Our attempts to fit the SO $_2$ lines with a variety of smooth models characterized by $T(0)$, T_N , and T_{atm} were not very successful. The observed shapes of the 222- and 143-GHz lines could only be matched at low atmospheric temperatures (below 150 K), in contrast to our previous models where the line profile could be scaled *ad libitum* by adjusting the atmospheric filling factors. At higher atmospheric temperatures, the synthetic lines are too narrow if the contrast is matched by adjusting $T(0)$ and T_N , as exemplified by Fig. 7. However, the low temperature solutions are not acceptable either because, due to the high pressures needed to account for the observed widths of the detected lines, the 146.605- and 146.550-GHz lines should have been detected. In addition, the associated SO $_2$ column densities are unreasonably large. For example, in the case shown in the short-dashed line on Fig. 7 ($T_{\text{atm}} = 136$ K, $T(0) = 135$ K, $T_N = 100$ K), the averaged SO $_2$ amount is 0.2 cm-am, much larger than the UV upper limits (see below). In spite of the uncertainties associated with the unknown horizontal variation of the atmospheric temperature, we tentatively conclude that the "smooth" vapor equilibrium models do not give an adequate description of the SO $_2$ distribution of Io's atmosphere.

3.6. Limits on Other Atmospheric Species

Our search for other atmospheric species was, in part, motivated by the idea, developed in the past, that the Pioneer 10 ionospheric densities require a noncondensable atmospheric component besides SO $_2$. This idea was supported by the preliminary analysis of our millimeter-wave data (Lellouch *et al.* 1990).

The data on the H $_2$ S, CO, and SO lines were used to set upper limits on the abundances of these gases, assuming various atmospheric temperatures. Unlike SO $_2$, H $_2$ S and CO do not condense on Io's dayside hemisphere, and their horizontal distribution was assumed to be uniform over Io's visible disk. Synthetic spectra in this case were generated by integration of the radiative transfer equation over the airmasses corresponding to Io's solid disk and adding the limb emission of the different atmospheric layers observed in horizontal viewing. The case for SO is somewhat different because it is the prime photodissociation product of SO $_2$. Therefore, SO was assumed to follow the same horizontal distribution as SO $_2$ and the upper limit was expressed as a SO/SO $_2$ mixing ratio.

Results are summarized in Table III. The upper limits on H $_2$ S and CO are of order 0.1 and 5 nbar, respectively. These values are valid for almost any range of atmospheric temperature, except at $T_{\text{atm}} \leq 170$ K, where less stringent upper limits apply. In this case, the limb emission dominates and results in a line which has a contrast above that due to the difference between the atmospheric and surface temperature. The typical upper limit on SO/SO $_2$ is 5% for temperatures exceeding 170 K. For lower temperatures, the emission in vertical viewing (airmass = 1) remains below the 3σ upper limit, whatever the SO abundance.

TABLE III
Upper Limits (3σ) on H₂S, SO, and CO

Temp. (K)	H ₂ S		CO		SO/SO ₂
	Pressure (bar)	Col. density (cm-am)	Pressure (bar)	Col. density (cm-am)	
130	3.5×10^{-9}	8×10^{-3}	1.4×10^{-8}	3×10^{-2}	No constraint
170	7×10^{-11}	1.5×10^{-4}	3×10^{-9}	6×10^{-3}	5%
300	6×10^{-11}	1.3×10^{-4}	3×10^{-9}	7×10^{-3}	6%
600	1.2×10^{-10}	2.7×10^{-4}	5×10^{-9}	1×10^{-2}	7%

4. DISCUSSION

4.1. Comparison with UV Data

In the past, there have been several attempts to measure the SO₂ atmospheric abundance from Io spectra obtained with the International Ultraviolet Explorer (IUE). Bertaux and Belton (1979) showed that the broad absorption near 2800 Å in a low-resolution (8 Å) albedo spectrum was consistent with a dense SO₂ atmosphere, although the contribution of SO₂ frost in this feature was difficult to estimate. Butterworth *et al.* (1980) used another low-resolution IUE spectrum to set an upper limit on the mean SO₂ abundance of 0.008 cm-am. This limit was, however, questioned by Belton (1982) who pointed out possible curve-of-growth effects at low temperatures, which would result in a too stringent upper limit. Recently, Ballester *et al.* (1990) analyzed high-resolution (0.2 Å) spectra of Io between 2900 and 3170 Å. Although these spectra did not show unambiguous SO₂ features, they could be used to place a refined upper limit on the SO₂ global abundance. Ballester *et al.* compared their data with a number of models for Io's SO₂ distribution, including (i) the locally buffered models with a surface temperature distribution proposed by Kumar (1985) and Summers (1985), (ii) the regional cold trapping models (Fanale *et al.* 1982), (iii) subsurface cold trapping models (Matson and Nash 1983), and (iv) insolation-propagation models (Brown and Matson 1987). Some sublimation models used a global SO₂ frost coverage, while others used a partial frost coverage and local control of the atmospheric pressure, yielding a patchy SO₂ atmosphere. They also investigated the effect of volcanic sources. While the signal-to-noise ratio of the data did not allow them to discard the collisionally thin models, the intermediate models ((ii) and (iv)) seemed to best match the spectra (see Figs. 5 and 8 of Ballester *et al.*). In fact, assuming a homogeneous SO₂ atmosphere, Ballester *et al.* obtained an upper limit on the column density of 2×10^{17} cm⁻² (0.0074 cm-am). Their analysis assumed that UV absorption by SO₂ scales linearly with the gas column density, an assumption which can be made if the absorption cross section corresponds to spectral

averages over unsaturated lines. The validity of this hypothesis is difficult to assess, since it depends on the unknown distribution of individual lines. Belton (1982) has shown that using a continuum model (the Beer-Lambert law) for SO₂ absorption instead of a model based on a line spectrum may lead to considerable errors in the computation of planetary spectra, due to line saturation. However, high atmospheric temperatures, such as those that we find, tend to relieve the saturation of the individual lines and to increase the "effective number" of lines involved in the absorption, possibly making valid the Beer-Lambert formulation used by Ballester *et al.* (1990). In this case, their upper limit of 0.0074 cm-am hardly depends on the exact distribution of SO₂ gas and may be almost directly compared with the global (disk-averaged) SO₂ densities as measured by us, with a correction due to the observing geometry. Ballester *et al.* integrated the SO₂ absorption over the entire disk, which is approximately equivalent to using an airmass of 4 (G. Ballester, private communication). We used an airmass of 1 for the millimeter-wave spectrum, corresponding to 2 for the solar reflected spectrum. Therefore, the 0.0074 cm-am value must be multiplied by a factor of 2 for comparison.

As shown in Fig. 5c, the resulting upper limit (0.015 cm-am) is easily consistent with the millimeter-wave data. Although a detailed comparison is difficult due to the uncertainty on the saturation of the UV lines, it also appears that the UV upper limit allows, to some extent, a reduction of the acceptable (T_{atm} , N) domains obtained from our fits of the detected lines. Specifically, the cold atmosphere models ($T_{\text{atm}} \leq 200$ K) are rejected independently of the constraint obtained from the nondetection of the 146-GHz lines.

4.2. Reassessment of the Voyager/IRIS Observation

The Voyager infrared experiment IRIS detected gaseous SO₂ in absorption at 7 μm in a hot spot near Loki (Pearl *et al.* 1979). The average temperature of the hot spot in the field-of-view, as indicated by the continuum brightness temperature, was approximately 205 K. Using a simple radiative transfer model, in which the atmosphere

was described as a single layer of temperature 130 K, Pearl *et al.* inferred a local pressure of about 10^{-7} bar, corresponding to a SO_2 column density of 0.2 cm-am. The region of Io observed by IRIS contains three features: (i) a plume, identified as Plume 2 in Smith *et al.* (1979), (ii) a warm region at 290 K covering 10% of the field of view, and (iii) the remaining area at about 130 K. As discussed in some detail by Pearl *et al.* (1979), the first two of these three sources are unlikely to produce an important SO_2 atmosphere; noting that a pressure of 10^{-7} bar corresponds to saturation at 130 K, Pearl *et al.* proposed that the observed absorption suggests that SO_2 gas is in equilibrium with the surface temperature. This conclusion has been questioned, however, because the absorption may have been caused by a transient flow of SO_2 gas from the plume (Collins 1981) and because the surface temperature measured by IRIS was 127 ± 2 K (Hanel *et al.* 1979).

An obvious question is the mutual consistency of the global atmospheric (pressure, temperature) conditions derived from our measurements and the local conditions as obtained from the IRIS data. In particular, while Pearl *et al.* did not attempt to derive any constraint on the gas temperature and assumed a value of 130 K, it is necessary to investigate whether the high temperatures inferred in the previous section are acceptable in view of the IRIS observation. If not, the case for a warm average atmosphere would appear very unlikely, as it does not seem reasonable that the atmosphere would be colder near the hot spots than on average.

The first remark regarding the Pearl *et al.* (1979) analysis is that the measured column density depends very much on the assumed atmospheric temperature, not only through the line intensities, but mostly because this band is formed in strongly saturated conditions (as is obvious from the moderate growth of the absorption from 0.1 to 1 cm-am in Fig. 2 of Pearl *et al.*). The degree of saturation of the lines is a function of the linewidth, which is entirely ruled by the temperature in a pure Doppler regime. If the typical atmospheric temperature is higher than assumed, the lines will tend to desaturate and significantly lower amounts of SO_2 are required to account for the observed absorption.

As compared to the continuum emission at 205 K, the self-emission of SO_2 is negligible at 130 K and the SO_2 band appears in absorption. At higher temperatures, however, the SO_2 emission may become significant, eventually overwhelming the absorption. But, the fact that SO_2 is observed in absorption against the 205-K background does not necessarily imply that the atmospheric temperature is less than 205 K, because the 7- μm emission occurs largely in NLTE conditions, and a complete analysis of the feature observed by IRIS requires a NLTE radiative transfer model.

A rigorous analysis of the NLTE radiative transfer

equation requires simultaneous solving of a set of equations which are coupled because the source function depends monochromatically on the outgoing intensity itself. Such a treatment has been performed recently in the case of the upper atmosphere of Titan by Yelle (1991). Instead, we have used the simpler formulation described by Houghton (1986), in which the NLTE source function can be expressed as (combining his Eqs. (5.32) and (5.37))

$$J_\nu = \frac{B_\nu}{1 + 5\tau_\nu^*/(12\Phi)},$$

where B_ν is the Planck function, Φ is the ratio of radiative to collisional lifetimes, and τ_ν^* is the probability of an emitted photon getting out to space:

$$\tau_\nu^* = \int d\nu k_\nu \exp\left(-\int \frac{5}{3} k_\nu(z) N(z) dz\right) / \int d\nu k_\nu,$$

where k_ν is the absorption coefficient at frequency ν . $\Phi = k_d N/A$, where A is the Einstein coefficient for the ν_3 band, N is the atmospheric concentration, and k_d is the deactivation rate in $\text{cm}^3 \text{sec}^{-1}$.

Vibrational relaxation in SO_2 has been measured for several temperatures by Bass *et al.* (1971). SO_2 has three fundamental bands: ν_2 (518 cm^{-1}), ν_1 (1151 cm^{-1}), and ν_3 (1361 cm^{-1}). However, at room and higher temperatures, the ν_1 and ν_3 bands are collisionally coupled and both modes relax in parallel. Bass *et al.* (1971) measured a relaxation rate of $6.8 \times 10^{-14} \text{ cm}^3 \text{sec}^{-1}$ at 298 K, 2.2×10^{-13} at 528 K, and 6.1×10^{-13} at 901 K. Since relaxation measurements are often uncertain, we have used these values as "nominal" numbers and have studied the impact of variations in these numbers on the synthetic spectra. Using $A = 41 \text{ sec}^{-1}$ at 300 K, and one of the fit models from the millimeter-wave observations ($T_{\text{atm}} = 300 \text{ K}$, $N = 2 \times 10^{11} \text{ cm}^{-3}$), we get $\Phi = 3 \times 10^{-4}$ at the ground level, indicating strong NLTE effects. However, the lines are very narrow and saturated ($\tau \sim 30$ in the core of the strongest) and optical depth effects may be very important.

The probabilities of photon escape, the source functions, and the resulting synthetic spectra were calculated in a multilayer atmospheric model using two different formalisms. One is a conventional line-by-line model, which monochromatically computes the absorption coefficients over the entire extent of the band. This procedure is the most accurate but is in practice very time consuming because the step of the computation must be less than the Doppler width of the lines (about $1 \times 10^{-3} \text{ cm}^{-1}$ at 300 K), which requires hundreds of thousands of points. However, the number of lines involved in the SO_2 emission is large enough that a statistical band model can be used

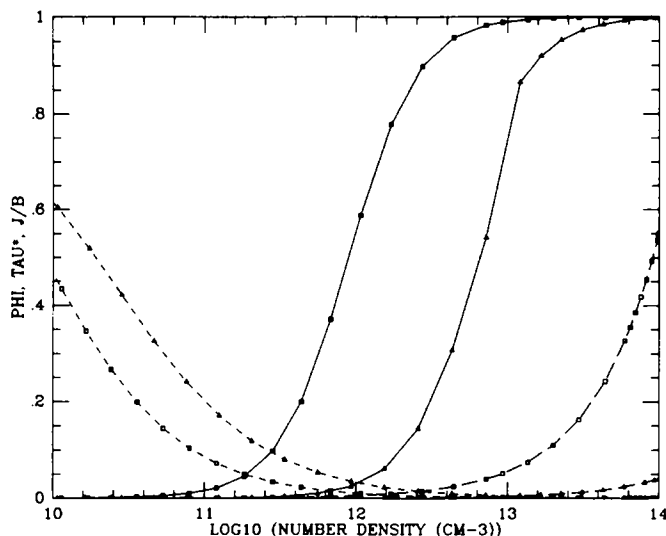


FIG. 8. Source function for the ν_3 band of SO_2 at 1360 cm^{-1} in Io's conditions. Long-dashed lines, ratio of radiative to collisional lifetime (Φ). Short-dashed lines, probability of atmospheric escape τ_v^* . Solid lines, ratio of source function to Planck function J_v/B_v . These three numbers are computed as a function of number density at temperatures of 130 K (triangles) and 500 K (squares).

even at the relatively fine resolution of the IRIS spectrum (4.3 cm^{-1}). The average transmission in each resolution element and the total equivalent width of the band were computed assuming an inverse line strength distribution of the line intensities and the formalism described by Wallace *et al.* (1974). In the framework of the band model, the escape probability can be simply computed by noting that $\tau_v^* = 1/S \, dW/du$, where $S = \int k_\nu \, d\nu$ is the integrated band intensity and W is the equivalent width of the band for integrated path length u . The two models were run using spectroscopic data from the GEISA bank (Husson *et al.* 1986) and were shown to be in agreement within several percent.

Figure 8 shows the values of Φ , τ_v^* , and the ratio J_v/B_v as a function of number density for two Io atmospheric models at temperatures of 130 and 500 K. The role of trapping is evidenced, lowering the density level where $J_v/B_v = 0.5$ (which one may define as the level of LTE breakdown) by almost two orders of magnitude. At Io's surface conditions ($N = 2 \times 10^{11} \text{ cm}^{-3}$), all the ν_3 emission occurs in NLTE conditions. However, this NLTE emission may well influence the spectrum, since it occurs at a much higher temperature than the surface.

Figure 9 shows the synthetic spectrum of the ν_3 band for an atmospheric temperature of 500 K and different SO_2 number densities, assuming the "nominal" relaxation rates of Bass *et al.* (1971). The spectrum of the background was calculated by taking the (wavenumber-dependent) background temperature from the IRIS spectrum. At densities larger than $1 \times 10^{11} \text{ cm}^{-3}$, the NLTE emission

overcomes the absorption, and the band appears in emission. At lower densities, the departure from LTE becomes so important that the ν_3 band appears in absorption. However, the band depth, which reaches a maximum near $N = 5 \times 10^{10} \text{ cm}^{-3}$ and eventually vanishes at lower densities, is always less than that observed, so that our baseline model cannot reproduce the IRIS spectrum at 500 K, and a fortiori at higher temperatures. To get agreement with the IRIS spectrum at $T_{\text{atm}} = 500 \text{ K}$, we must arbitrarily divide the relaxation rates by a factor of 5; the number density is then $9 \times 10^{10} \text{ cm}^{-3}$ (Fig. 10).

At all temperatures below 400 K, the IRIS spectrum can be fitted with our NLTE emission baseline model. Examples at 130 K and 300 K are shown in Fig. 10, where the corresponding densities have been determined in order to reproduce the observed equivalent width. At 130 K, we find $N = 2 \times 10^{12} \text{ cm}^{-3}$ (i.e., $p = 4 \times 10^{-8}$ bar, column density = 0.09 cm-am). These values are approximately 2.5 times lower than those given by Pearl *et al.* (1979). Since the absorption varies slowly with absorber amount because of saturation, it seems reasonable to attribute the difference to the use of different SO_2 absorption coefficients and to the fact that Pearl *et al.* used a single homogeneous layer model. At temperatures where the emission is significant, the observed spectrum can be fitted with two distinct densities, because of the effect of the pressure on the NLTE emission; the modeled band is too deep between the two values, and too shallow or in emission outside. At 300 K, we find $N = 1.6 \times 10^{11} \text{ cm}^{-3}$ (0.016 cm-am , 6 nbar) and $N = 8 \times 10^{11} \text{ cm}^{-3}$ (0.07 cm-am , 32 nbar). We note that the first of these cases is

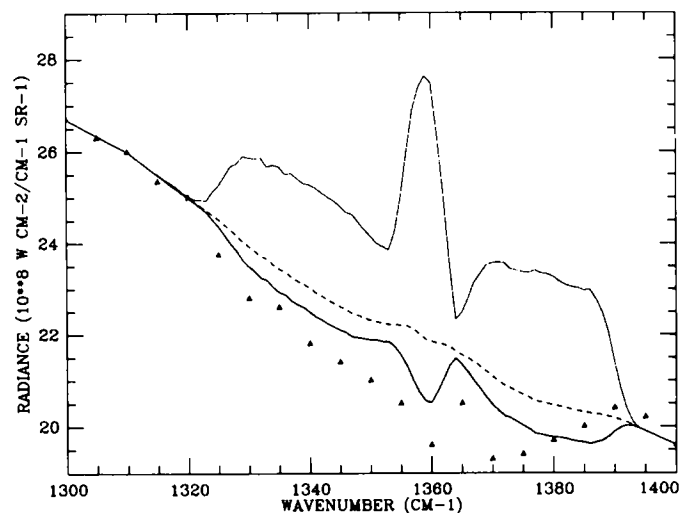


FIG. 9. Synthetic spectrum in the ν_3 band of SO_2 at $T_{\text{atm}} = 500 \text{ K}$, compared to the IRIS/Voyager spectrum of Loki (triangles). Solid thick line, $N = 5 \times 10^{10} \text{ cm}^{-3}$. Dashed line, $N = 10^{11} \text{ cm}^{-3}$. Solid thin line, $N = 2 \times 10^{11} \text{ cm}^{-3}$.

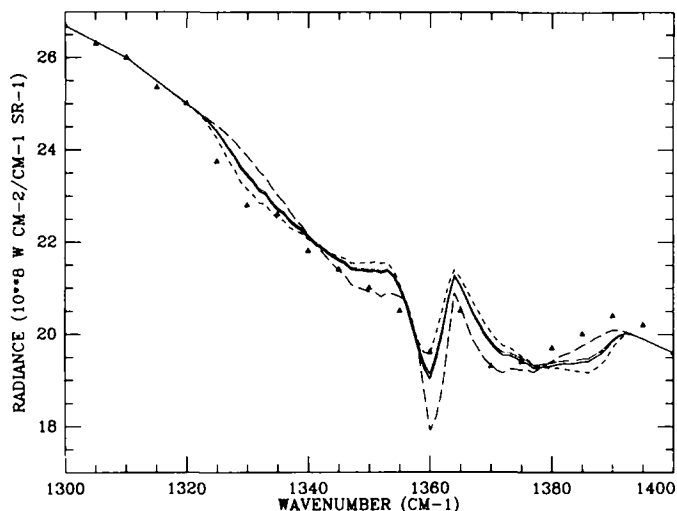


FIG. 10. Best fit of the Loki spectrum (triangles) for several atmospheric temperatures. Long dashed line: $T_{\text{atm}} = 130 \text{ K}$, $N = 2 \times 10^{12} \text{ cm}^{-3}$. Solid line: $T_{\text{atm}} = 300 \text{ K}$, $N = 1.6 \times 10^{11} \text{ cm}^{-3}$. Dotted line: $T_{\text{atm}} = 300 \text{ K}$, $N = 8 \times 10^{11} \text{ cm}^{-3}$. Short-dashed line: $T_{\text{atm}} = 500 \text{ K}$, $N = 9 \times 10^{10} \text{ cm}^{-3}$. In this last case, the relaxation coefficient of the (ν_1, ν_3) system of SO_2 has been arbitrarily divided by 5 compared to the value given in Bass *et al.* (1971).

very close to the preferred solution found by fitting the 143-, 222-, and 346-GHz lines. Although (Fig. 10) the IRIS spectrum cannot be fitted in every detail, it is remarkable that the shape is better reproduced at 300 K (and in the “modified” 500 K case, where the relaxation coefficient is divided by 5) than at 130 K—at which temperature the Q branch is too strong compared to the wings. Assuming that the observed absorption is entirely due to the ν_3 band of gaseous SO_2 , we take it as an additional indication of a high temperature for Io’s atmosphere, at least near Loki.

In summary, the Loki observation can be fitted for a variety of atmospheric temperatures (up to at least 400 K and possibly more if our crude treatment of the NLTE case overestimates the emission). The corresponding range of surface pressure varies from about 5 to 40 nbar, corresponding to vapor pressure equilibrium temperatures of 119–126 K. Better agreement is obtained if the atmosphere is heated to high temperatures (300 K or more), although very high temperatures (above 500 K) seem excluded because our model puts the ν_3 band in emission rather than in absorption. At this point, it is worth reiterating that our treatment of the NLTE case is simplified and that more detailed models are necessary to reach a definite conclusion. Comparing these results with the pressure and temperature derived from the millimeter-wave data, we take the striking similarity as a strong indication that the atmosphere we globally observe from Earth is effectively the same as that seen locally by Voyager. Whether the atmospheric distribution is determined

by sublimation equilibrium or emission from volcanic sources cannot be decided on the basis of this analysis alone, however, because the conditions at Loki when the spectrum was taken (a hot spot in the equatorial region at 13 hr) allow both types of atmospheres to be significant.

The question of why gaseous SO_2 has been detected by IRIS only near Loki has been raised in the past. The common answer is that over most of Io’s disk, the surface intercepted by the IRIS field-of-view is not warm enough to be detected at $7 \mu\text{m}$, thus a fortiori an absorption band. (There are, however, a few exceptions (e.g., Pele, Amirani/Maui . . .) and the uniqueness of the detection of SO_2 at Loki, if related to volcanic activity, may indicate that other plumes are not driven by SO_2 , or contain much less SO_2 gas than the Loki plume (Pearl and Sinton 1982). This explanation is satisfactory if the gas is cooler than the surface background, but it must be examined whether it remains valid for a warm atmosphere, since in this case one may expect the ν_3 band to be detectable in emission even though the surface is undetectable. To investigate this question, we have computed synthetic spectra of this band for the various atmospheric cases discussed above, assuming a surface temperature of 120 K (corresponding to a continuum of $0.025 \times 10^{-8} \text{ W cm}^{-1} \text{ sr}^{-1}$). If $T_{\text{atm}} = 130 \text{ K}$, $N = 2 \times 10^{12} \text{ cm}^{-3}$, the ν_3 band is in absorption. At $T_{\text{atm}} = 300 \text{ K}$, the band is in emission, and the maximum flux (at 1360 cm^{-1}) is $0.15 \times 10^{-8} \text{ W cm}^{-1} \text{ sr}^{-1}$ for $N = 1.6 \times 10^{11} \text{ cm}^{-3}$ and $2.5 \times 10^{-8} \text{ W cm}^{-1} \text{ sr}^{-1}$ for $N = 8 \times 10^{11} \text{ cm}^{-3}$. Finally, in the “modified” 500 K case ($N = 9 \times 10^{10} \text{ cm}^{-3}$ and the relaxation coefficient is divided by 5), the maximum flux in the band is $0.4 \times 10^{-8} \text{ W cm}^{-1} \text{ sr}^{-1}$. The typical sensitivity of the IRIS spectra at Io is approximately $0.5 \times 10^{-8} \text{ W cm}^{-1} \text{ sr}^{-1}$ (147 K at 1360 cm^{-1}). We conclude that in all previous situations, except in the case $T_{\text{atm}} = 300 \text{ K}$, $N = 8 \times 10^{11} \text{ cm}^{-3}$, the band would have been undetectable by IRIS. Taken with the fact that the best model obtained by fitting only the detected millimeter-wave lines corresponds to $T_{\text{atm}} = 300 \text{ K}$, $N = 2 \times 10^{11} \text{ cm}^{-3}$, this suggests that $N = 1.6 \times 10^{11} \text{ cm}^{-3}$ may be the more appropriate solution for $T_{\text{atm}} = 300 \text{ K}$. More importantly, this leaves open the possibility that gaseous SO_2 in the same pressure and temperature conditions as those observed at Loki were present over a number of other, cooler regions observed by IRIS.

4.3. Thermal Models

Kumar (1980, 1985) has computed a 1-D thermal structure model that included heating by solar UV and charged particles (electrons), cooling by NLTE infrared emission in SO_2 vibrational bands, and conductive heat transport. He simultaneously constructed ionospheric models and used the Pioneer 10 profiles to constrain the magnitude and distribution of the precipitating electron flux. The

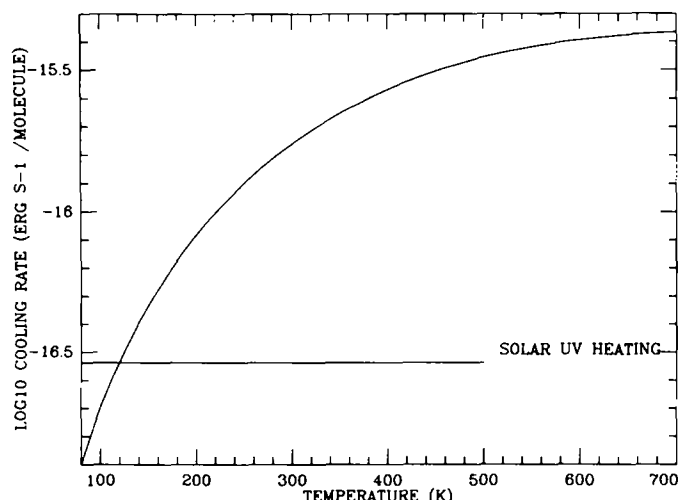


FIG. 11. The rotational cooling rate of SO_2 as a function of temperature, compared with the UV solar heating.

neutral atmosphere that corresponds to his best ionospheric model has a rapidly increasing temperature immediately above the surface, rising to about 400 K at 20 km and 800 K at 400 km.

While Kumar's models are elaborate in the sense that they properly address the coupling between photochemistry and thermal structure, his treatment of the radiative cooling can be improved in two respects. First, the laboratory work by Bass *et al.* (1971) on SO_2 deactivation coefficients, which shows that SO_2 does not behave like CO_2 , should be used. Second, the radiative cooling due to rotational transitions of SO_2 must be included. Omitting these two aspects leads to a severe underestimate of the cooling.

4.3.1. Rotational cooling by SO_2 . Yelle (1991), in his model of Titan's upper atmosphere, has shown how important the role of rotational cooling in molecular bands can be in the heat budget. Applying his approach to the case of Io, it can be shown that the balance between solar UV heating and SO_2 rotational cooling alone leads to atmospheric temperatures much lower than those derived from our observations.

Solar UV radiation heats the atmosphere by dissociation and ionization of SO_2 shortward of 2320 Å. The UV heating rate at solar zenith angle θ is

$$H = \int \epsilon \Phi_{\text{sol}}(\lambda) hc / \lambda N \sigma(\lambda) e^{-\tau(\lambda)} d\lambda,$$

where λ is wavelength, $\Phi_{\text{sol}}(\lambda)$ is the solar flux at 5.2 AU, $\sigma(\lambda)$ is the absorption cross section, and ϵ is the heating efficiency. $\tau(\lambda) = N \sigma(\lambda) H_a \chi(\theta)$ is the opacity at wavelength λ , where N is the SO_2 density (cm^{-3}), H_a is the scale height, and χ is Chapman's function.

In the optically thin regions of the atmosphere, the UV heating rate per molecule reduces to

$$H = \int \epsilon \Phi_{\text{sol}}(\lambda) hc / \lambda \sigma(\lambda) d\lambda.$$

We used solar fluxes for average conditions (Torr and Torr 1985) and cross sections from Golomb *et al.* (1962), Thompson *et al.* (1963), and Wu and Judge (1981). We assumed, as a rough estimate, a constant heating efficiency, $\epsilon = 0.35$. More refined calculations of ϵ would require a detailed study of ionospheric chemistry, well beyond the scope of this paper. With these values, we obtain $H = 2.9 \times 10^{-17} \text{ erg sec}^{-1}/\text{molecule}$. The corresponding photodissociation rate is $8.0 \times 10^{-6} \text{ sec}^{-1}$, in agreement with Kumar (1980).

The rotational cooling occurs in LTE in most of the atmosphere and can be expressed as

$$R = 4\pi N \int B_\nu k_\nu e^{-\tau_\nu} d\nu$$

where B_ν is the Planck function and k_ν and τ_ν are the absorption coefficient and opacity at frequency ν . In the optically thin limit, the cooling rate becomes $R = 4\pi \sum_j B_\nu S_j$ per molecule, where S_j is the strength of the j th rotational line.

R varies with temperature through the Planck function and the intensity line distribution. Balance of R and H in their optically thin limit is reached for an exospheric temperature of 120 K (Fig. 11). Consideration of the heat transfer equation in spherical geometry,

$$-\frac{1}{r^2} \frac{d}{dr} \left(k(T) \frac{dT}{dr} \right) = H - R,$$

yields the profile of Fig. 12, with a maximum temperature of 205 K near 20 km and an isothermal exospheric temperature at 120 K above 100 km. This profile is far too cool to account for our observations.

This simple calculation omits two factors. One is the breakdown of LTE for the rotational cooling at high altitudes, which will lead to higher exospheric temperatures. The other is the NLTE cooling in the vibrational bands, which scales as the square of the pressure, and will importantly modify the lower atmosphere thermal structure, possibly leading to a mesopause in the first scale height (Lellouch *et al.* 1990). More detailed computations (D. Strobel, private communication) confirm that atmospheric temperatures of 600 K at 40 km cannot be reproduced by heat transfer models including solar UV heating and SO_2 cooling only.

4.3.2. Plasma ion heating. Another source of energy

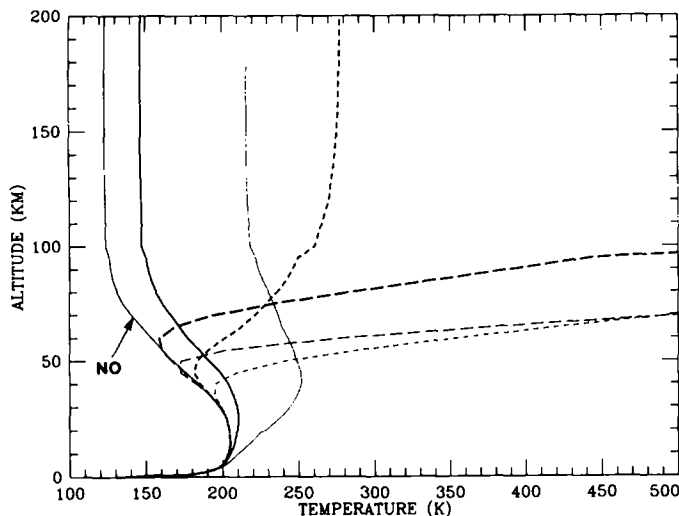


FIG. 12. Several computed thermal profiles for Io's lower atmosphere. The surface number density is assumed to be $8 \times 10^{11} \text{ cm}^{-3}$. Solid line marked "NO", profile obtained by considering only solar UV heating and rotational LTE cooling. Other profiles include additional heating by a plasma flux of $4 \times 10^{12} \text{ eV cm}^{-2} \text{ sec}^{-1}$ (Johnson *et al.* 1989). Thick lines correspond to 20% heating efficiency and thin lines to 100% efficiency. Several cross-sections for plasma interaction are considered: solid, 10^{-17} cm^2 ; short-dashed, 10^{-16} cm^2 ; long-dashed, 10^{-15} cm^2 .

is ion heating by magnetospheric plasma. Its effect on the atmospheric temperature was studied by Johnson (1989), who assumed deposition of energy in a thin layer corresponding to penetration of a SO_2 column density of $N_p = 6 \times 10^{15} \text{ cm}^2$, and heating by 20% of the average energy flux from the torus ($\Phi_p = 4 \times 10^{12} \text{ eV cm}^{-2} \text{ sec}^{-1}$). Neglecting other sources and energy sinks, which he assumed would only modify the lower atmosphere thermal structure, he analytically solved the heat transfer equation and obtained temperatures of thousands of degrees at altitudes of only a few tens of kilometers above the surface of Io.

We have adapted his approach to include plasma heating in our numerical estimates for the thermal structure. Plasma heating was assumed to occur with a cross section σ_p and a global efficiency ϵ_p . This description, which formally treats the particle heating as photon interaction, has been used earlier by Friedson and Yung (1984) in their modeling of electron impact on Titan's thermosphere; σ_p and the effective energy flux $\Phi_{\text{eff}} = \epsilon_p \Phi_p$ are free parameters of the model. Starting with parameters equivalent to those used by Johnson (1989), i.e., $\sigma_p = 10^{-16} \text{ cm}^2$ and $\Phi_{\text{eff}} = 1.3 \text{ erg cm}^{-2} \text{ sec}^{-1}$, and including the previous solar heating and the rotational cooling rate, we obtain an exospheric temperature of 280 K (Fig. 12). Much higher exospheric temperatures can be obtained if larger values of Φ_{eff} are assumed, but the temperatures below 40 km are unchanged from the case without particle heating. If

the cross section is higher than 10^{-16} cm^2 , the atmosphere is more readily heated up, but only at higher altitudes. If the cross section is lower, then the lower atmosphere is warmed, but the process is not very efficient, so that huge particle bombardment is necessary to achieve warm temperatures. Specifically, for a cross section of 10^{-17} cm^2 , even if a plasma energy flux of $4 \times 10^{12} \text{ eV cm}^{-2} \text{ sec}^{-1}$ (100% efficiency) is used, the model produces maximum temperatures of 250 K (Fig. 12).

Conversely, we can estimate the conditions of the plasma heating that are needed to produce the hot lower atmosphere implied by the millimeter observations, say 600 K at 40 km. With a surface number density of $2 \times 10^{11} \text{ cm}^{-3}$, the column density above 40 km is about $4 \times 10^{17} \text{ cm}^{-2}$. Particle heating down to this altitude would therefore imply a cross section of only $2.5 \times 10^{-18} \text{ cm}^2$. In addition, to get high temperatures, the vertically integrated plasma heating should be comparable to the integrated SO_2 rotational cooling at 600 K ($\sim 60 \text{ erg cm}^{-2} \text{ sec}^{-1}$, i.e., $4 \times 10^{13} \text{ eV cm}^{-2} \text{ sec}^{-1}$), since the solar heating is much less ($\sim 3 \text{ erg cm}^{-2} \text{ sec}^{-1}$) than the latter. Both of these conditions seem unlikely. Considering also the dynamical heat losses due to winds generated by horizontal pressure gradients, we conclude that another heat source is required.

4.4. SO_2 Equilibrium or Volcanic Atmosphere?

At this point, we can attempt to put all of our results together to try to infer the most likely nature of the SO_2 atmosphere of Io. A firm conclusion is that the dayside atmosphere does have a collisionally thick component, with a typical pressure of 10 nbars. Thin atmosphere models, such as the pure sputtering models and subsurface cold trap models, are readily excluded. It is more difficult to distinguish between the limiting cases of "equilibrium sublimation" atmosphere and an atmosphere in dynamical equilibrium with an array of volcanic sources. In favor of the former are the following arguments: (i) the atmosphere did not show any appreciable temporal variability between January 1990 and January 1991, (ii) the domain of SO_2 partial pressure and horizontal extent corresponds to vapor pressure equilibrium with surface frosts for the range of surface temperatures expected at the subsolar regions on Io. However, if the lower atmosphere is truly hot (500 K), "smooth" vapor equilibrium models (e.g., Fanale *et al.* 1982) are excluded. More likely is a patchy distribution of frosts (McEwen *et al.* 1988), controlled as in the hydrodynamical model by Ingersoll (1989), in which each region, of linear dimension $L = \sqrt{2\pi H/\alpha}$ (where H is the atmospheric scale height and α the sticking coefficient), controls its own pressure. Assuming $\alpha = 0.5$, $L = 180 \text{ km}$ for $T_{\text{atm}} = 500 \text{ K}$, an atmosphere covering approximately 5% of Io's visible disk (i.e., $5.2 \times 10^5 \text{ km}^2$) would require approximately 16 regions of frost deposits,

which seems a viable situation. However, in order to maintain their long-term stability, these patches require embedded volcanic sources (Kumar 1979, Ingersoll 1989).

In the framework of equilibrium atmosphere models, the difference we observe between the two sides of Io would have to correspond to a difference in surface pressure (or, equivalently, the temperature of the frost) rather than in coverage (i.e., extent or number of subliming frost patches). This seems somewhat surprising because there are more extensive SO₂ frost deposits on the leading than on the trailing side (Howell *et al.* 1984, McEwen *et al.* 1988). The factor-of-four pressure enhancement on the trailing side (14 nbar vs 3 nbar on the leading) translates into a 5-K difference (122 K vs 117 K) in the effective temperature of the frosts. If the sublimation model is correct, our observations therefore imply that the frosts are colder on the leading side.

The "volcanic atmosphere" hypothesis may provide a better explanation for Io's atmosphere. Voyager 2 detected nine active plumes, of which eight were still active by the time of Voyager 1, 4 months later (Strom and Schneider 1982). Volcanoes are capable of generating substantial (if transitory) atmospheres over large spatial scales (Lee and Thomas 1980). Ingersoll (1989) finds that a flux of $3.5 \times 10^{-6} \text{ kg m}^{-2} \text{ sec}^{-1}$, equivalent to a 10^6 kg/sec source, produces an atmospheric pressure of about 22 nbar over a 300 km radius. This is in the pressure range of what we have observed. It is worth noting that the major volcanic plumes observed by Voyager (Pele and Loki) are on the trailing side. This may be why we observe the higher pressure there. Regarding the temperature of the plumes, Pearl and Sinton (1982) have made a detailed analysis of the hot spots observed by IRIS/Voyager and concluded that three types of sources can be distinguished: (i) stable low temperature ($\leq 400 \text{ K}$) sources, (ii) transient high temperature (600 K) sources, (iii) stable high temperature (600 K) sources. Groundbased eclipse measurements by Sinton (1980) showed evidence for a nearly constant flux from sources near 560 K over a period of 9 days, with high temperature features being probably small and numerous. A large hot source (654 K) was also observed by Voyager to be associated with Pele. Therefore, it seems possible that the SO₂ emission we observe from Earth is due to a combination of numerous hot sources, which on average produce a hot atmosphere in dynamical equilibrium. Theoretical estimates of the temperature of the plume gas are, however, extremely uncertain (Kieffer 1982).

We therefore feel that the volcanic atmosphere concept provides the most adequate representation of Io's atmosphere, although sublimation equilibrium probably plays a role as well. Related to the nature of Io's atmosphere is the issue of its location. Sublimation atmospheres are located where the frosts are, and volcanic atmospheres

are centered around volcanoes. However, since the frosts are supplied from the volcanoes, they must be where the volcanic sources are. Indeed, the SO₂ frost is concentrated within $\pm 30^\circ$ of latitude (McEwen *et al.* 1988), which is also where most of the active plumes were observed. From this, it can be inferred that most of the SO₂ atmosphere takes place between $\pm 30^\circ$. Regarding the latitudinal distribution of the atmosphere, the natural place for the sublimation equilibrium component is around the warmest areas, in particular in the subsolar regions. However, assuming that the temperature of the frost does not depend on its geographical location, but only on its latitude and solar zenith angle, an atmosphere entirely confined to the subsolar regions would lead to similar pressures on both sides, with possibly a larger atmospheric coverage of the surface on the leading side, due to the larger frost coverage. This contradicts the observation and points to the existence of an "off-centered" atmospheric component, uncorrelated with the SZA.

4.5. Upper Limits on the Global Composition of the Atmosphere

Our upper limits on H₂S, SO, and CO are useful in constraining the global composition of Io's atmosphere; however, we emphasize that they are contingent to our conclusion on the atmospheric temperature (Table II).

SO is the prime photodissociation product of SO₂ and, according to current aeronomical models (e.g., Kumar 1985), is expected to be present at the 10% level of SO₂. Although SO is undoubtedly present, we find an upper limit of 5% of SO₂, which should be considered by future photochemical models.

The H₂S case is more intriguing. A detection of H₂S in the solid phase at $3.9 \mu\text{m}$ has been claimed by Nash and Howell (1989) and Salama *et al.* (1990). H₂S is thought to be a time-variable component of Io's surface, intimately mixed with SO₂. Gaseous H₂S has also been proposed as a major component of Io's atmosphere (Matson *et al.* 1988) on the basis of the following arguments: (i) the S/O ratio in the magnetospheric torus exceeds the stoichiometric SO₂ value, i.e., another source of S is required, (ii) SO₂ alone cannot account for the ionospheric densities observed by Pioneer 10 (see below), (iii) H₂S is a common volcanic gas along with SO₂, and (iv) a spectrally distinct, dark, polar cap unit has been discovered on Io (McEwen *et al.* 1988), whose spectral reflectance is consistent with that of H₂S frost darkened by exposure to UV radiation (Lebofsky *et al.* 1976).

The lack of laboratory measurements on the vapor/solid equilibrium of H₂S in the relevant temperature range makes the conversion of our upper limit of 10^{-10} bar to equivalent equilibrium temperature very uncertain. If, nonetheless, we use data measured at 134 and 115 K by

Kraus *et al.* (1989) to extrapolate at lower temperatures (using $P(\text{atm}) = 5.8 \times 10^{**}(7-1479/T(\text{K}))$), we find a limiting temperature of 83 K, which is on the lower end of the range of temperatures generally accepted for the high latitudes on Io (Johnson and Matson 1989). (A temperature of 72 K is obtained if we use the vapor pressure formula quoted by Moreno *et al.* (1991), namely $P(\text{atm}) = 2550 \exp(-2217/T(\text{K}))$.) Our upper limit, therefore, may still allow the possibility that H_2S polar caps buffer the global abundance of gaseous H_2S , but excludes that gaseous H_2S be in equilibrium with solid in other, warmer places. If the identification of H_2S at $3.9 \mu\text{m}$ is correct, this is a puzzling result because H_2S is much more volatile than SO_2 , and might be expected to be the uppermost unit and therefore totally sublimed. Salama *et al.* (1990) counter this by arguing that trapping of highly diluted (3%) H_2S in a SO_2 -dominant mixture could raise its sublimation temperature, and allow it to remain in the solid state even at Io's observed surface temperature. However, this does not solve the problem because the H_2S equilibrium vapor pressure in the mixture ($P_{\text{H}_2\text{S},\text{mix}}$) is approximately given by (B. Schmitt, private communication):

$$P_{\text{H}_2\text{S},\text{mix}}/P_{\text{H}_2\text{S},\text{pure}} = f[\text{SO}_2/\text{H}_2\text{S}]_{\text{ice}},$$

where $f = 1$ –10. Even if one takes $f = 1$, the H_2S equilibrium pressure is lowered by a factor 30 at most, and our upper limit (10^{-10} bar) corresponds to only 91 K. The detection of solid H_2S is therefore a serious difficulty. We note that laboratory studies (Nash *et al.* 1986, de Bergh *et al.* 1991) show that thick deposits of SO_2 exhibit features at $3.9 \mu\text{m}$, which makes the H_2S identification questionable.

CO has not been considered a strong candidate for the neutral atmosphere because the Io torus does not contain any C-bearing species. Interest for carbon molecules, nonetheless, has been restored by the tentative identification of CO_2 clusters at $2.12 \mu\text{m}$ (Trafton *et al.* 1990, Sandford *et al.* 1991). This identification, however, has been recently questioned (deBergh *et al.* 1991). Our upper limit on CO, about 3 nbar, excludes CO as a major component of Io's atmosphere, as least in the regions where SO_2 is present.

4.6. The Pioneer 10 Ionospheric Layers

Despite the complications associated with the Pioneer 10 observation (Johnson and Matson 1989) it has been argued that a neutral atmosphere of at least several nanobars in the vicinity of the terminator was required to support the ionosphere densities measured by Pioneer 10 (e.g., Kumar 1985). If SO_2 is in strict local vapor equilibrium, the surface pressure at the terminator is unlikely to

exceed about 4×10^{-12} bar (equilibrium at 100 K), therefore SO_2 cannot account for the ionosphere. Our upper limit on H_2S also rules it out as the source of stable ionospheric layers. Molecular oxygen, a noncondensable photolysis product of the photolysis of SO_2 , has been advocated by Kumar and Hunten (1982) as a plausible global atmospheric constituent. A simple vertical model of Io's thermal budget (Lellouch *et al.* 1990), which considers solar heating as the only energy input, shows that an O_2 pressure of 5–40 nbar is required to produce exospheric temperatures consistent with those inferred from the ionospheric scale heights measured by Pioneer 10. This was our main argument for proposing that Io may have a background atmosphere of about 20 nbar. CO, with a maximum pressure of 3 nbar, is unable to sustain such high temperatures, because of intense LTE cooling by its rotational transitions.

The need for another atmospheric species to explain the Pioneer 10 ionosphere has been challenged by Ingersoll (1989), who proposed that the observed ionospheric layers could be the result of a local atmosphere maintained by an active volcanic vent. Supporting this hypothesis is the fact that the Pioneer 10 egress site is near Loki and the ingress site has bright SO_2 deposits, which presumably must be maintained against sublimation losses by volcanic activity. Ingersoll shows that the propagation of gas emitted from a source that abruptly terminates onto a sufficiently cold surface will lead to an horizontal atmospheric distribution with a decay length determined by the atmospheric temperature. The distance between the Pioneer 10 egress point (17 N, 291 W) and Loki (19 N, 305 W) is about 450 km. That between the ingress point (15 S, 111 W) and the closest hot spot, Prometheus (3 S, 153 W), is about 1400 km. Ingersoll's (1989) analysis (his formula (A.29)), with an atmospheric temperature of 500 K ($L = 180$ km), yields local pressures at egress and ingress sites equal to 1/5 of that at Loki and 1/30 of that at Prometheus, respectively. These pressures are probably large enough to account for the ionosphere and provide a natural explanation of the Pioneer 10 occultation in terms of SO_2 . If the atmospheric temperature in the first scale height is low, e.g., 125 K ($L = 45$ km), these numbers become 1/70 and 10^{-5} , and at least for the ingress ionosphere, this model no longer works. Therefore, the high temperatures in the lower atmosphere apparently implied by our observations may be the key to explaining the Pioneer 10 ionospheric layers. This explanation can account for different ionospheric densities and peak altitude at ingress and egress, as observed (Kliore *et al.* 1974), although such differences may as well be due to variations in the strength of the ionizing source (Cloutier *et al.* 1978).

5. SUMMARY

The main conclusions of this study can be summarized as follows:

(1) Millimeter-wave observations show that Io has a detectable SO₂ neutral atmosphere. This atmosphere is, at least partly, time-independent.

(2) The SO₂ atmosphere has a typical partial pressure of 10 nbar (column density: 0.02 cm-am, or 6×10^{17} cm⁻²) and covers a reduced fraction of Io's surface (5–20%). The SO₂ pressure seems larger (by a factor ~4) on the trailing side than on the leading side.

(3) If the atmosphere is assumed not to show any time variability, observations of SO₂ lines with different intensities indicate that the lower atmosphere must be quite hot: 500–600 K at 40 km altitude.

(4) Models in vapor pressure equilibrium with a smooth distribution of surface frosts do not seem to give an adequate description of the horizontal distribution of atmospheric SO₂.

(5) Non-LTE radiative transfer models show that the Voyager/IRIS observation of SO₂ at Loki, originally analyzed for an atmospheric temperature of 130 K, can actually be fitted for a variety of temperatures (up to at least 400 K). High temperatures give a better agreement to the data, and the corresponding SO₂ local pressures (5–40 nbar) are in striking coincidence with those inferred from millimeter-wave data.

(6) Simple thermal models, including heating by solar UV and particle precipitation, and radiative cooling by SO₂, seem unable to reproduce high temperatures in the lower atmosphere.

(7) Although the apparent stability and the measured SO₂ column densities are strong arguments in favor of equilibrium vapor pressure models, we believe that the most accurate picture of the atmosphere is provided by dynamical models with volcanic sources, where relatively dense local atmospheres are generated by volcanoes over spatial lengths proportional to the vertical scale height.

(8) No other atmospheric constituent (H₂S, SO, CO) was detected by millimeter-wave observations. The absence of H₂S is disturbing in view of its claimed detection in the solid phase.

(9) The hot (~500 K) lower atmosphere originating in volcanic sources that is indicated by our observations provides a natural explanation of the Pioneer 10 occultation measurements of ionospheric layers which does not require the assumption of an additional component or a special geographical situation.

APPENDIX: SUMMATION OF INDIVIDUAL IO SCANS

As explained in Section 2, the spectra obtained at IRAM had to be corrected for the velocity of Io relative to Jupiter. This is done by resampling individual spectra to a common grid before averaging. Each original spectrum is sampled on a grid defined by the central frequencies, ν_i , and channel spacing $\Delta\nu$. The frequency response of channel i is denoted by $R_i(\nu)$. The shape of the filter response for each channel is very similar and we can assume that $R_i(\nu) = R_{i-1}(\nu + \Delta\nu) = R(\nu - \nu_i)$,

where R is the average shape of the channel response. The grid of a resampled spectrum is defined as a set ν'_j of frequencies with the same channel spacing as the original spectrum. The temperature T'_j of channel j of a resampled spectrum is estimated as the weighted average of the two adjacent channels in the original spectrum:

$$T'_j = (1 - \alpha)T_i + \alpha T_{i+1} \quad \text{where } \nu_i \leq \nu'_j < \nu_{i+1} \text{ and } \alpha = (\nu'_j - \nu_i)/\Delta\nu.$$

Over the time it takes to obtain a meaningful signal-to-noise ratio, the frequency shift of an Io line is much more than the spacing between two channels (except at 3 mm with the 1-MHz filterbanks where it spans only two channels), and therefore it can be assumed with reasonable precision that α is uniformly distributed between 0 and 1. Under these conditions, a simple integration shows that the equivalent filter bandshape of the sum of the resampled spectra is the convolution product of the original filter bandshape by a triangle function whose base width is twice the channel spacing,

$$R'(\nu - \nu'_j) = \int_{\nu - \Delta\nu}^{\nu + \Delta\nu} \frac{d\nu'}{\Delta\nu} \left[1 - \frac{|\nu' - \nu|}{\Delta\nu} \right] R(\nu' - \nu'_j).$$

This effect has to be taken into account because we observe lines only a few channels wide with sharp edges characteristic of saturation, which are smoothed by the resampling process.

ACKNOWLEDGMENTS

We are grateful to the IRAM and CSO staffs for their assistance during the observations, and to W. Thuillot, J.-E. Arlot, J. Lieske, and M. Standish for providing us with the Io ephemeris. We also thank G. Ballester and D. M. Hunten for interesting comments. I.d.P. acknowledges support from NSF Grant AST-8900156. M.B. acknowledges the support of NASA Grant W-17469. The National Optical Astronomy Observatories are operated by Associated Universities, Inc., under contract with the NSF.

REFERENCES

- BAARS, J. W. M., B. G. HOOGHOUT, P. G. MEZGER, AND M. J. DE JONG 1987. The IRAM 30-M millimeter radio telescope on Pico Veleta, Spain. *Astron. Astrophys.* **175**, 319–326.
- BALLESTER, G. E., D. F. STROBEL, H. W. MOOS, AND P. D. FELDMAN 1990. The atmospheric abundance of SO₂ on Io. *Icarus* **88**, 1–23.
- BARTHOLDI, P., AND F. OWEN 1972. The occultation of Beta Scorpii by Jupiter and Io. II. Io. *Astron. J.* **77**, 60–65.
- BASS, H. E., T. G. WINTER, AND L. B. EVANS 1971. Vibrational relaxation in SO₂. *J. Chem. Phys.* **54**, 644–647.
- BAUMGARDNER, J., M. A. MORENO, G. SCHUBERT, AND M. G. KIVELSON 1987. Two classes of volcanic eruptions and their corresponding atmospheres on Io. *Bull. Am. Astron. Soc.* **19**, 856.
- BELTON, M. J. S. 1982. An interpretation of the near-ultraviolet absorption spectrum of SO₂. Implications for Venus, Io, and laboratory measurements. *Icarus* **52**, 149–165.
- BERTAUX, J. L., AND M. J. S. BELTON 1979. Evidence of SO₂ on Io from UV observations. *Nature* **282**, 813–815.
- BROWN, R. H. 1973. In *Exploration of the Planetary System: Proc. of the Copernicus Symposium 4* (IAU Symposium 65). (A. Woszczyk and C. Iwaniszewska, Eds.).
- BROWN, R. H., AND D. L. MATSON 1987. Thermal effects of insolation propagation into the regolith of airless bodies. *Icarus* **72**, 84–94.
- BUTTERWORTH, P. S., J. CALDWELL, V. MOORE, T. OWEN, A. R.

- RIVOLO, AND A. L. LANE 1980. An upper limit to the global SO₂ abundance on Io. *Nature* **285**, 308–309.
- CHENG, A. F., P. K. HAFF, R. E. JOHNSON, AND L. J. LANZEROTTI 1986. Interactions of planetary magnetospheres with icy satellites surfaces. In *Satellites* (J. A. Burns and M. S. Matthews, Eds.), pp. 403–436. Univ. of Arizona Press, Tucson.
- CLOUTIER, P. A., R. E. DANIELL, JR., A. J. DESSLER, AND T. W. HILL 1978. A cometary ionosphere model for Io. *Astrophys. Space Sci.* **55**, 93–112.
- COLLINS, S. A. 1981. Spatial color variations in the volcanic plume at Loki, on Io. *J. Geophys. Res.* **86**, 8621–8626.
- DEBERGH, C., E. LELLOUCH, B. SCHMITT, AND J.-P. MAILLARD 1991. New FTS observations of Io at 2 and 4 microns. *Bull. Am. Astron. Soc.* **23**, 130.
- ELLISON, B. N., P. L. SCHAFFER, W. D. VAIL, AND R. E. MILLER 1989. A 345 GHz SIS receiver for radio-astronomy. *Int. J. IR MM Waves* **10**, 937–948.
- FANALE, F. P., R. H. BROWN, D. P. CRUIKSHANK, AND R. N. CLARK 1979. Significance of absorption features in Io's IR reflectance spectrum. *Nature* **280**, 760–761.
- FANALE, F. P., W. BANERDT, L. ELSON, T. V. JOHNSON, AND R. ZUREK 1982. In *Satellites of Jupiter* (D. Morrison, Ed.), pp. 756–781. Univ. of Arizona Press, Tucson.
- FRIEDSON, A. J., AND Y. L. YUNG 1984. The thermosphere of Titan. *J. Geophys. Res.* **89**(A1), 85–90.
- GERARD, E. 1987. The OH radio lines in comets: A review. In *Cometary Radio Astronomy, Green Bank, WV: NRAO Workshop 17*. (W. M. Irvine, F. P. Schloerb, and L. Tacconi-Garman Eds.), p. 91.
- GOLOMB, D., K. WATANABE, AND F. F. MARMO 1962. Absorption coefficients of sulfur dioxide in the vacuum ultraviolet. *J. Chem. Phys.* **36**, 958–961.
- HANEL, R. A., B. CONRATH, M. FLASAR, V. KUNDE, P. LOWMAN, W. MAGUIRE, J. PEARL, J. PIRRAGLIA, R. SAMUELSON, D. GAUTIER, P. GIERASCH, L. HORN, S. KUMAR AND C. PONNAMPERUMA 1979. Infrared observations of the Jovian system from Voyager 1. *Science* **204**, 972–976.
- HOUGHTON, J. T. 1986. *The Physics of Atmospheres*, pp. 61, 65. Cambridge Univ. Press, London/New York.
- HOWELL, R. R., D. P. CRUIKSHANK, AND F. P. FANALE 1984. Sulfur dioxide on Io: Spatial distribution and physical state. *Icarus* **57**, 53–92.
- HUSSON, N., A. CHEDIN, N. A. SCOTT, D. BAILLY, G. GRANER, N. LACOME, A. LEVY, C. ROSSETTI, G. TARRAGO, C. CAMY-PEYRET, L. M. FLAUD, A. BAUER, J. M. COLMONT, N. MONNANTEUIL, J. C. HILICO, G. PIERRE, M. LOETE, J. P. CHAMPION, L. S. ROTHMAN, L. R. BROWN, G. ORTON, P. VARANASI, C. P. RINSLAND, M. A. H. SMITH, AND A. GOLDMAN 1986. The GEISA spectroscopic line parameters data bank in 1984. *Ann. Geophys.* **4**, 185–190.
- INGERSOLL, A. P., M. E. SUMMERS, AND S. G. SCHLIPF 1985. Supersonic meteorology of Io: Sublimation driven flow of SO₂. *Icarus* **64**, 375–390.
- INGERSOLL, A. P. 1989. Io meteorology: How atmospheric pressure is controlled locally by volcanos and surface frosts. *Icarus* **81**, 298–313.
- JOHNSON, R. E. 1989. Plasma heating of an SO₂ atmosphere on Io. *Geophys. Res. Lett.* **16**, 1117–1120.
- JOHNSON, T. V., AND D. L. MATSON 1989. Io's tenuous atmosphere. In *Origin and Evolution of Planetary and Satellites Atmospheres* (S. K. Atreya, J. B. Pollack, and M. S. Matthews, Eds.), pp. 661–682. Univ. of Arizona Press, Tucson.
- KIEFFER, S. W. 1982. Dynamics and thermodynamics of volcanic eruption: Implications for the plumes on Io. In *Satellites of Jupiter* (D. Morrison, Ed.), pp. 647–723. Univ. of Arizona Press, Tucson.
- KLIORE, A. J., D. L. CAIN, G. FJELDBO, B. L. SIEDEL, M. SYKES, AND S. I. RASOOL 1974. Preliminary results on the atmospheres of Io and Jupiter from Pioneer 10 S-band occultation experiment. *Science* **183**, 323–324.
- KLIORE, A. J., G. FJELDBO, B. L. SIEDEL, D. N. SWEETNAM, T. T. SESPLAUKIS, AND P. M. WOICESHYN 1975. Atmosphere of Io from Pioneer 10 occultation measurements. *Icarus* **24**, 407–419.
- KRAUS, G. F., J. E. ALLEN, JR., AND L. C. COOK 1989. Vapor pressure measurements of hydrogen sulfide. *Bull. Am. Astron. Soc.* **21**, 948.
- KUMAR, S. 1979. The stability of an SO₂ atmosphere on Io. *Nature* **280**, 758–760.
- KUMAR, S. 1980. A model for the atmosphere and ionosphere of Io. *Geophys. Res. Lett.* **7**, 9–12.
- KUMAR, S. 1982. Photochemistry of SO₂ in the atmosphere of Io and implications on atmospheric escape. *J. Geophys. Res.* **87**, 1677–1693.
- KUMAR, S. 1985. The SO₂ atmosphere and ionosphere of Io: Ion chemistry, atmospheric escape, and models corresponding to the Pioneer 10 radio occultation measurements. *Icarus* **61**, 101–123.
- KUMAR, S., AND D. M. HUNTEN 1982. The atmospheres of Io and other satellites. In *Satellites of Jupiter* (D. Morrison, Ed.), pp. 647–723. Univ. of Arizona Press, Tucson.
- LANZEROTTI, L. J., AND W. L. BROWN 1983. Supply of SO₂ for the atmosphere of Io. *J. Geophys. Res.* **88**, 989–990.
- LEBOFSKY, L., AND M. B. FEGLEY, JR. 1976. Laboratory reflection spectra for the determination of chemical composition of icy bodies. *Icarus* **28**, 379–387.
- LEE, S. W., AND P. C. THOMAS 1980. Near-surface flow of volcanic gases on Io. *Icarus* **44**, 280–290.
- LELLOUCH, E., M. BELTON, I. DE PATER, S. GULKIS, AND T. ENCRENAZ 1990. Io's atmosphere from microwave detection of SO₂. *Nature* **346**, 639–641.
- MATSON, D. L., AND D. NASH 1983. Io's atmosphere: Pressure control by regolith cold trapping and surface venting. *J. Geophys. Res.* **88**, 4771–4783.
- MATSON, D. L., T. V. JOHNSON, A. S. MCEWEN, AND L. SODERBLOM 1988. Io: The case for an H₂S atmosphere. *Bull. Am. Astron. Soc.* **20**, 1125.
- MCGRATH, M., AND R. E. JOHNSON 1987. Magnetospheric plasma sputtering of Io's atmosphere. *Icarus* **69**, 519–531.
- MCEWEN, A. S., T. V. JOHNSON, D. L. MATSON, AND L. A. SODERBLOM 1988. The global distribution, abundance, and stability of SO₂ on Io. *Icarus* **75**, 450–478.
- MORENO, M. A., G. SCHUBERT, J. BAUMGARDNER, M. G. KIVELSON, AND D. A. PAIGE 1991. Io's volcanic and sublimation atmospheres. *Icarus* **93**, 63–81.
- NASH, D. B., M. CARR, J. GRADIE, D. M. HUNTEN, AND C. YODER 1986. Io. In *Satellites* (J. A. Burns and M. S. Matthews, Eds.), pp. 629–688. Univ. of Arizona Press, Tucson.
- NASH, D. B., AND R. R. HOWELL 1989. Hydrogen sulfide: Evidence from telescopic and laboratory infrared spectra. *Science* **244**, 454–457.
- PEARL, J. C., R. HANLE, V. KUNDE, W. MAGUIRE, K. FOX, S. GUPTA, C. PONNAMPERUMA, AND F. RAULIN 1979. Identification of gaseous SO₂ and new upper limits for other gases on Io. *Nature* **288**, 757–758.
- PEARL, J. C., AND W. M. SINTON 1982. Hot spots on Io. In *Satellites of Jupiter* (D. Morrison, Ed.), pp. 724–755. Univ. of Arizona Press, Tucson.
- POYNTER, R. L., AND H. M. PICKETT 1982. *JPL Publ. 80-23 Rev. 1*. Jet Propulsion Laboratory, Pasadena.
- SALAMA, F., L. J. ALLAMANDOIA, F. C. WITTEBORN, D. P. CRUIKSHANK, S. A. SANDFORD, AND J. D. BREGMAN 1990. The 2.5–5.0 μ m

- spectra of Io: Evidence for H_2S and H_2O frozen in SO_2 . *Icarus* **83**, 66–82.
- SANDFORD, S. A., F. SALAMA, L. J. ALLAMANDOLA, L. M. TRAFTON, D. F. LESTER, AND T. F. RAMSEYER 1991. Laboratory studies of the newly discovered infrared band at 4705.2 cm^{-1} ($2.1253\text{ }\mu\text{m}$) in the spectrum of Io: The tentative identification of CO_2 . *Icarus* **91**, 125–144.
- SINTON, W., A. TOKUNAGA, E. BECKLIN, I. GATLEY, T. LEE, AND C. J. LONSDALE 1980. Io: Ground-based observations of hot spots. *Science* **210**, 1015–1017.
- SIEVEKA, E. M., AND R. E. JOHNSON 1985. Non-isotropic coronal atmosphere on Io. *J. Geophys. Res.* **90**, 5327–5331. Correction: **91**, 4608.
- SMITH, B. A., AND S. A. SMITH 1972. Upper limits for an atmosphere on Io. *Icarus* **17**, 218–222.
- SMITH, B. A., L. A. SODERBLUM, T. V. JOHNSON, A. P. INGERSOLL, S. A. COLLINS, E. M. SCHOEMAKER, G. E. HUNT, H. MASURSKY, M. H. CARR, M. E. DAVIES, A. R. COOK II, J. BOYCE, G. E. DANIELSON, T. OWEN, C. SAGAN, R. F. BEEBE, J. VEVERKA, R. G. STROM, J. F. MCCAULEY, D. MORRISON, G. A. BRIGGS AND V. E. SUOMI 1979. The Jupiter system through the eyes of Voyager 1. *Science* **204**, 951–972.
- SMYTHE, W. D., R. M. NELSON, AND D. B. NASH 1979. Spectral evidence for SO_2 frost or absorbate on Io's surface. *Nature* **280**, 766–767.
- SNYDER, L. E., P. PALMER, AND I. DE PATER 1989. Radio Detection of Formaldehyde Emission from Comet Halley. *Astron. J.* **97**, 246–253.
- STROM, R. G., AND N. M. SCHNEIDER 1982. Volcanic eruption plumes on Io. In *Satellites of Jupiter* (D. Morrison, Ed.), pp. 598–633. Univ. of Arizona Press, Tucson.
- SUMMERS, M. E. 1985. *Theoretical Studies of Io's Atmosphere*. Ph.D. thesis, California Institute of Technology, Pasadena.
- THOMPSON, B. A., P. HARTECK, AND R. R. REEVES, JR. 1963. Ultraviolet absorption coefficients of CO_2 , CO , O_2 , H_2O , NH_3 , NO , SO_2 , and CH_4 between 850 and 4000 Å. *J. Geophys. Res.* **68**, 6431.
- TORR, M. G., AND D. G. TORR 1985. Ionization frequencies for solar cycle 21: Revised. *J. Geophys. Res.* **90**, 6675–6678.
- TRAFTON, L. M., D. F. LESTER, T. F. RAMSEYER, F. SALAMA, S. A. SANDFORD, AND L. J. ALLAMANDOLA 1991. A new class of absorption feature in Io's near-infrared spectrum. *Icarus* **89**, 264–276.
- WAGMAN, D. D. 1979. *Sublimation Pressure and Enthalpy of SO_2* . Thermodynamics Data Center, Natl. Bur. Standards, Washington D.C.
- WALLACE, L., M. PRATHER, AND M. J. S. BELTON 1974. The thermal structure of the atmosphere of Jupiter. *Astrophys. J.* **193**, 481–493.
- WU, C. Y. T., AND D. L. JUDGE 1981. Study of sulfur-containing molecules in the EUV region. I. Photoabsorption cross section of SO_2 . *J. Chem. Phys.* **74**, 3804–3806.
- YELLE, R. V. 1991. Non-LTE models of Titan's upper atmosphere. *Astrophys. J.* **383**, 380–400.

Article V

FIRST ABSOLUTE WIND MEASUREMENTS IN THE MIDDLE ATMOSPHERE OF MARS

EMMANUEL LELLOUCH,¹ JEFFREY J. GOLDSTEIN,² STEPHEN W. BOUGHER,³
 GABRIEL PAUBERT,⁴ AND JAN ROSENQVIST¹

Received 1991 March 11; accepted 1991 June 11

ABSTRACT

The first absolute wind measurements in the middle atmosphere of Mars (40–70 km) were obtained from Doppler shifts in the $J = 2-1$ CO transition at 230.538 GHz. During the 1988 opposition, this line was observed at 100 kHz resolution with the IRAM 30 m telescope. The 12" FWHM beam of the facility allowed spatial resolution of the Martian disk (23"8). The high S/N of the data allowed measurement of winds with a 1σ absolute line-of-sight accuracy of 20 m s^{-1} . The measurements, performed during Southern summer solstice ($L_s = 279^\circ$), stress the Southern hemisphere and clearly indicate a global easterlies flow. If modeled by a broad easterly jet with a maximum centered at 20 S, and extending 80° in latitude, the jet core velocity is found to have a χ^2 minimum at 160 m s^{-1} ($\pm 80 \text{ m s}^{-1}$ at the 95% confidence level), generally consistent with predictions for broad summer easterly jets near 50 km as proposed by theoretical models. If the flow is modeled instead by a planet-wide solid rotator zonal flow which is restricted to the Southern hemisphere or equatorial regions, the velocity of the easterlies is nearly the same. These wind measurements, together with the temperature measurements of Deming et al., provide the first experimental rough picture of the middle atmosphere circulation of Mars, in general agreement with the Jaquin axisymmetric middle atmosphere model and the current Mars GCM model of Pollack et al.

Subject headings: planets: atmospheres — planets: Mars — radio sources: spectra

1. INTRODUCTION

While our knowledge of the dynamics of the lower atmosphere of Mars (0–50 km) has benefitted from extensive temperature measurements by the *Mariner 9* IRIS (Hanel et al. 1972) and *Viking* IRTM (Martin 1981), and from direct surface wind measurements by the *Viking* landers (Hess et al. 1977), an understanding of the circulation at higher altitudes (50–100 km) is presently limited by lack of observational constraints. The middle atmosphere circulation of Mars is believed to be quite similar to that of the Earth. On Earth, solstitial differential heating drives a mean meridional circulation with rising motion near the summer pole, a meridional drift at high levels into the winter hemisphere, and sinking motion near the winter polar mesopause. The Coriolis torques exerted on this meridional flow generate mean zonal easterlies (retrograde) in the summer hemisphere, and westerlies (prograde) in the winter hemisphere (Leovy 1964; Holton 1975). A strong zonal friction force is additionally required to satisfy both the heat and momentum budgets of the terrestrial mesosphere (Lindzen 1981; Holton 1982, 1983; Holton & Zhu 1984). The mechanism producing such drag is believed to be the breaking, near the turbopause, of gravity waves generated in the Earth's troposphere by flow over topography, etc. This momentum stress (drag) due to breaking gravity waves appears to be primarily responsible for "closing off" the mesospheric jets at high altitude (both the westerly winter and the easterly summer), and for the reversal of the latitudinal temperature gradient in the

upper mesosphere. Temperatures are thus lowest at the summer pole (due to adiabatic cooling from upwelling) and warmest at the winter pole (due to adiabatic heating from subsidence). In view of the large amplitude of Martian topography, gravity waves should play a role at least as important for the atmosphere of Mars as for that of the Earth.

Observational data related to the middle atmosphere dynamics of Mars are limited to temperature maps from heterodyne measurements of CO₂ 10 μm lines performed during the 1984 opposition (Deming et al. 1986; Rothermel et al. 1988). The temperatures, averaged over the 50–85 km altitude range were found to vary greatly with latitude, with minimum temperatures (close to CO₂ radiative equilibrium) near the subsolar point, and strong warming (by 50–100 K over CO₂ radiative equilibrium temperatures) towards high latitudes in both the winter and summer hemispheres. As in the terrestrial case, the reversed winter hemisphere latitudinal temperature gradient is likely driven by dynamical heating in the winter polar region (Deming et al. 1986). The latitudinal symmetry of the observed temperatures is possibly due to increased radiative heating, especially at high summer latitudes, due to absorption by suspended dust (Haberle, Leovy & Pollack 1982). These measurements have been extremely useful to constrain dynamical models of the middle atmosphere of Mars (Jaquin 1989; Barnes 1990). However, they do not provide a straightforward picture of its wind system.

We report here the first absolute wind measurements in the middle atmosphere of Mars. They are based on inspection of Doppler shifts of the $J = 2-1$ ¹²CO rotational line, mapped over the Martian disk during the 1988 opposition. A similar study has been successfully achieved for Venus by Pierce, Muhleman, & Berge (1991) (in $J = 1-0$ CO) and Buhl, Goldstein, & Chin (1991) (in $J = 2-1$ CO). The objective of our observations was in fact the spatial variability of Martian carbon monoxide abundance which was obtained from simul-

¹ Département de Recherches Spatiales, Observatoire de Paris-Meudon, 92195 Meudon, France.

² Laboratory for Astrophysics, National Air and Space Museum, Smithsonian Institution, Washington, DC 20560.

³ Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721.

⁴ Institut de Radio-Astronomie Millimétrique, Avda de la Divina Pastora, Granada, Spain.

taneous observations of the $J = 2-1$ ^{12}CO and ^{13}CO lines (Lellouch et al. 1991). However, the signal-to-noise ratio on the ^{12}CO 2-1 line appeared to be high enough to unambiguously detect Doppler shifts and to allow wind measurements with a typical 1σ absolute accuracy of 20 m s^{-1} .

2. OBSERVATIONS

On 1988 September 23–25, we observed Mars in the $J = 2-1$ ^{12}CO and ^{13}CO rotational lines at 230 and 220 GHz, respectively, using the IRAM 30 m radiotelescope at Pico Veleta, Spain. The telescope was equipped with a SIS receiver characterized by a receiver noise temperature of about 250 K. Sky conditions were excellent (zenith opacity less than 0.1), allowing system temperatures of 500–600 K. Taking advantage of the narrow beam of the facility [about $12''$ full width at half-maximum (FWHM) at 230 GHz], compared to the angular diameter of Mars ($23.8''$), we mapped the lines as a function of the position on the disk. Fourteen points were recorded, roughly corresponding to a sampling at half-beam. Since our observation was performed almost exactly at opposition, perihelion, and Southern summer solstice ($L_s = 279^\circ$), the entire dayside of Mars was observed, spatial resolution was maximized and extensive coverage was achieved for the Southern hemisphere. The sampled latitudes extend from 60° N to 90° S , with five points recorded along the equator. The directions east and west mentioned throughout are planetary based. The ^{12}CO and ^{13}CO spectra were obtained by rapid position switching with the telescope secondary wobbling at 1 Hz. Observations were performed in double sideband, with the line located in the lower sideband. We used two filter banks, a 512 channel 1 MHz resolution bank and a 256 channel 100 kHz resolution bank. At 230 GHz, 100 kHz resolution corresponds to a line-of-sight (LOS) velocity resolution of 130 m s^{-1} . However, the signal-to-noise (S/N) allowed line centroiding to $\sim 15\text{ kHz}$ (1σ) or 20 m s^{-1} .

The ^{12}CO $J = 2-1$ line at Mars is seen in absorption with a 25% depth and 20 MHz width, whereas the ^{13}CO line generally appears as a 3% absorption feature of 1 MHz width superposed on broader emission wings due to the lower atmosphere (nadir and limb) contribution. Because our first objective was to measure the CO abundance, for which the ^{13}CO line is the best tracer, most of the observing time was dedicated to achieve a satisfactory S/N on ^{13}CO . Our 100 kHz resolution observations of the ^{12}CO line, which contain the Martian winds information, are in fact limited to two 1 minute spectra for each of the 14 points. As detailed below, the probed altitudes extend from about 40 to 70 km. Although sufficient for an accurate retrieval of the CO abundance, the S/N on the ^{13}CO lines is not large enough to infer winds at lower altitudes.

To minimize pointing errors, which are critical in this kind of observation, the observations were generally performed as follows. The pointing was checked by azimuth and elevation continuum drifts across Mars. Then after a calibration procedure was performed, 1 minute integration scans were recorded successively at the 14 target points. Pointing and calibration were then rechecked, followed by a second set of scans at each point. The 100 kHz resolution observations of the ^{12}CO line were obtained when Mars elevation was between 51° and 42° (air mass 1.29 to 1.49). The estimated absolute pointing uncertainty, from residuals in continuum drifts, is $3''$. Given the way the experiment was conducted, the pointing error is systematic i.e., the magnitude and direction of the offset are the same for

all beam locations. The relative (offset) pointing and tracking uncertainty are estimated to be about $0.5''$ each.

3. DATA REDUCTION AND MEASUREMENT OF BEAM-INTEGRATED WINDS

For each of the $^{12}\text{CO}(2-1)$ 100 kHz resolution spectra, a 1 MHz resolution spectrum was recorded simultaneously. This "low" resolution spectrum was used to define the continuum level. Baseline ripples, which are known to affect millimeter-wave spectra of planets, were eliminated by removing a first-degree baseline at low resolution, then applying the same baseline correction to the corresponding high resolution (100 kHz) spectrum. The baseline was always found flat over the 25.6 MHz bandwidth of the high-resolution spectrum.

In the Lellouch Paubert, & Encrenaz (1991) analysis, the individual ^{12}CO and ^{13}CO spectra were rescaled to their continuum levels, and then averaged by position offsets without any further treatment. This allowed us to infer, for each beam position, the local CO mixing ratio and vertical thermal structure. In particular the uncertainty on the retrieved thermal structure was about 20 K, primarily due to the uncertainty on the local surface temperature.

Accurate wind measurements from Doppler shifts in the $^{12}\text{CO}(2-1)$ line center requires fine frequency corrections to the data and improved modeling. The real-time software at IRAM, before each scan, computes the total velocity of the source with respect to the telescope (Earth-Mars translation plus Earth's rotation), and adjusts the local oscillator (LO) frequency accordingly. While the LO, in principle, refers the spectra to the barycentric frame of Mars, its step size is exactly 24.48 kHz, which means that the sky frequency cannot be synthesized with an absolute accuracy better than $\pm 12.24\text{ kHz}$. This limitation was overcome by recomputing Mars-IRAM relative velocity at the time of each scan, comparing the required LO frequency with its true value, and shifting the frequency scale of the scan by the appropriate fractional number of channels. Each pair of 100 kHz spectra for each beam position were then averaged.

The Doppler shifts were obtained by cross-correlating the observed averaged spectra within $\pm 1\text{ MHz}$ of line core (over 20 high-resolution channels) with synthetic profiles generated by a radiative transfer model, using the local CO and thermal structure measurements of Lellouch et al. (1991). Absolute wind velocity determination is limited by both the accuracy of the rest frequency for the 2-1 transition and the value of the pressure shift in CO by CO_2 . We adopted the recently measured values of $230538.000 \pm 0.001\text{ MHz}$ ($\pm 1.3\text{ m s}^{-1}$) and $+18\text{ kHz/torr}$ (at 180 K), respectively (S. P. Belov et al. 1991 private communication). Both were incorporated into the radiative model which was otherwise identical to that of Lellouch et al. (1991). The pressure shift at the level where the line core is formed [about $3\text{ }\mu\text{bar}$ (70 km)] is negligible. Superposition of the core onto pressure-shifted wings was found to induce a minor core shift of 1 kHz (1.3 m s^{-1}). Retrieved lineshifts are relatively insensitive to the exact shape of the reference profile (e.g., Pierce et al. 1991), therefore are very little affected by the radiative transfer model uncertainties. By successively measuring the shift at a given point of the map using synthetic profiles corresponding to various other points, we estimated this effect to be at the $1-2\text{ m s}^{-1}$ level.

In addition, fitting was also independently performed to the data (1) within $\pm 0.35\text{ MHz}$ of line core, and (2) in the $\pm(0.35-1$

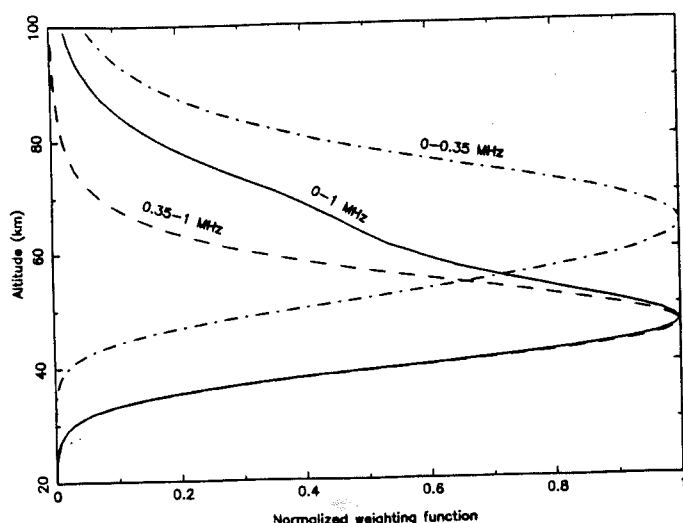


FIG. 1.—Average weighting functions in the $^{12}\text{CO } J = 2-1$ line at 230.538 GHz, computed for a CO mixing ratio of 8×10^{-4} and the observing geometry of the subearth point.

MHz) spectral region, in an attempt to derive vertical wind velocity information. In each of the three cases, the best model fit was obtained by shifting the synthetic spectrum along the observed by steps of 1 kHz, and the standard deviation, σ , of the data from this best model was computed. Random noise at the obtained σ level was then generated, superposed on the best model and the resulting distribution was refit. After 20 repetitions, the standard deviation in the offset from line rest frequency was computed. In addition to the spectroscopic uncertainties mentioned earlier, systematic errors due to the accuracy of the ephemeris ($\leq 1 \text{ m s}^{-1}$) and to a gravitational blueshift (1 m s^{-1}) are negligible.

In order to assign altitude levels for the wind measurements, the temperature weighting functions were computed for an average mixing ratio of 8×10^{-4} and the observing geometry at the subearth point. The weighting functions at 0, 0.35 and 1 MHz from line center peak at 71, 55, and 42 km, respectively, with half-power widths of 21, 18, and 13 km, respectively. Averaging the weighting functions between 0 and 1 MHz from

line center shows that the 0–1 MHz interval provides information on levels between 39 and 63 km altitude (the half-power altitudes for the averaged weighting functions). Similarly, the 0–0.35 and 0.35–1 MHz intervals probe the 52–78 km and 39–56 km ranges, respectively (Fig. 1). Although it would be of obvious interest to extend measurements further into the line wings, the flatness of the line shape beyond 1 MHz from line center prevents retrieval of meaningful shifts at lower altitudes. For example, the 1σ uncertainty on winds measured between 1 and 4 MHz from line center (probing the 25–42 km altitude range) would be typically 60 m s^{-1} .

4. RESULTS

4.1. Qualitative Analysis

Table 1 gives the *beam-integrated* Doppler shifts (expressed in m s^{-1} , positive for recession) measured over the three intervals: 0–1, 0–0.35, and 0.35–1 MHz from line center. The corresponding 1σ wind uncertainties are 21, 25, and 30 m s^{-1} on average. As examples, spectra recorded along the Martian equator are shown in Figure 2, along with their best model fits.

Figure 3 shows the LOS projected winds measured in the 0–1 MHz range, after correction for beam-integrated Mars solid planet rotational components, as detailed below. Immediately evident, particularly in the Southern hemisphere, is a general zonal easterlies flow characterized by approach in the east, near zero LOS velocities across the subearth meridian and recession in the west. Beam position 4 near the South pole, with a strong blueshift cannot be interpreted in terms of a purely zonal field and was therefore omitted from the modeling. Although in principle, comparison of points off the central meridian can reveal information about latitudinal variations, the large beam was found to preclude such studies here. Finally, inspection of Table 1 does not reveal any obvious variation of the circulation with altitude at the vertical resolution considered.

4.2. Dynamical Modeling

A retrieved wind velocity reflects dynamical contributions from all regions intercepted by the beam. Modeling of these *beam-integrated* wind velocities is therefore required for a quantitative determination of the global wind field responsible.

TABLE 1
OFFSET MEASUREMENTS^a

Point	Latitude	Longitude ^b	0–1 MHz	0–0.35 MHz	0.35–1 MHz
1.....	–21	0	–1(12)	16(13)	–18(13)
2.....	0	90	–60(26)	–20(17)	–97(59)
3.....	0	–90	87(27)	96(39)	81(31)
4.....	–90	0	–120(27)	–210(44)	–98(29)
5.....	60	0	–9(20)	–18(23)	–5(29)
6.....	0	0	–13(16)	–6(19)	–25(17)
7.....	0	35	–57(19)	–73(32)	–38(21)
8.....	0	–35	60(18)	61(25)	38(23)
9.....	–35	0	12(14)	35(21)	–13(25)
10.....	25	0	7(12)	16(13)	–25(19)
11.....	–35	45	–45(32)	–57(31)	–53(41)
12.....	–35	–45	5(27)	20(28)	–13(40)
13.....	25	60	–116(23)	–114(43)	–118(38)
14.....	25	–60	44(21)	73(20)	–21(28)

^a In m s^{-1} ; positive offsets indicate recession; brackets: 1σ uncertainty.

^b Positive planet westward from the central meridian.

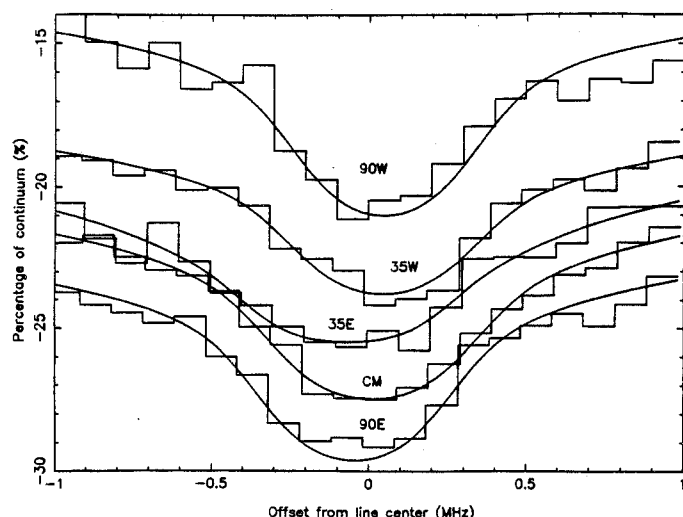


FIG. 2.—Line cores observed along the Martian equator (histograms) compared to best-fit models (solid lines). The line center is assumed to be at 230.538000 GHz. The line absorption is expressed in continuum units. “CM” indicates central meridian. Longitudes are measured from the CM.

Modeling of the retrieved winds was accomplished (as per Goldstein et al. 1991; Goldstein 1990) by assuming a global, zonal wind field superposed on the solid planet rotation of Mars (242 m s^{-1} at the equator) and convolving the Doppler shifted line contributions over the entire beam. The zonal field was initially assumed to be a solid body rotator (horizontal velocity is proportional to the cosine of the latitude). Vertical velocities were assumed everywhere zero. A model space was generated by modeling the beam-integrated winds for different

zonal wind strengths. Goodness-of-fit between the *ensemble* of modeled and observed velocities allows model wind field discrimination.

The beam was partitioned into a 100×100 element grid, with a beam diameter twice the diffraction-limited FWHM ($12''$). Within each beam element, horizontal wind velocity, LOS projection geometry, and beam spatial roll-off were taken as constant and evaluated at the element centroid. Each element produces a line contribution, frequency translated by the appropriate LOS velocity, as dictated by the model wind field. The contribution's relative weight is determined from the beam roll-off and element area. The element contributions were spectrally summed, and the frequency (velocity) centroid of the convolution was defined as the LOS beam-integrated velocity. Martian solid rotation components used to generate Figure 3 were obtained through beam modeling with the wind field removed.

Model goodness-of-fit was determined from the (error bar weighted) dispersion S between the N observed and modeled wind velocities, in units of the mean error bar (Goldstein 1990). A minimum S near unity indicates a good model fit with a dispersion attributable to the rms noise in the data. However, for small N , a model space about the minimum exists within which models are statistically indistinguishable. Quantitative description of this “best-fit model space” was obtained using the χ^2 test. A 95% confidence level best-fit model space was adopted, outside which the χ^2 probability of obtaining the observed S through statistical fluctuations is $\leq 5\%$. This required, for $N = 13$ retrieved wind velocities, that S be less than 1.3.

Regarding the ± 1 MHz data, a number of model spaces were generated with zonal wind fields of varying latitudinal (β) extent and parameterized by equatorial velocity V_{eq} . As detailed in Table 2, the observed data could only be adequately fit by a definitive easterly flow that must at least exist in tropical to mid Southern latitudes. For the planet-wide case ($-90^\circ \leq \beta \leq 90^\circ$), $V_{eq} = 110 \pm 55 \text{ m s}^{-1}$. The error reflects the extent of the 95% confidence level model space. Restricting the wind field to Southern latitudes or the tropics boosts the strength of V_{eq} by $50\text{--}75 \text{ m s}^{-1}$ (on average). Restricting the wind field to Northern latitudes yields a minimum $S = 1.4$, precluding a fit to the data.

Another characterization of the zonal wind field was attempted, more in line with current simulations of easterlies near 50 km from the Mars general circulation model (GCM) (Pollack et al. 1990) and a Mars climate model under development (J. R. Barnes 1991, private communication). See discussion below for details. A broad summer easterly jet with a sinusoidal rolloff was assumed, characterized by a maximum centered at 20°S and zeros at 20°N and 60°S . Only the jet's core velocity was allowed to vary to obtain a best-fit model. The model space minimum occurs for a jet core velocity at 20°S of $160 \pm 80 \text{ m/s}$, much the same as the restricted Southern latitude (case 2) and tropical (case 4) zonal flow models above (see Table 2).

The nominal beam positions were then offset by the $3''$ worst case pointing uncertainty East, West, North, and South on the planet, to determine the effects on the best-fit model spaces. East and West offsets produced model spaces unable to fit the data: beam-integrated velocities of $20\text{--}30 \text{ m s}^{-1}$ were obtained at positions along the subearth meridian, inconsistent with the observed (Fig. 3). Pointing uncertainty if indeed present at the $3''$ level, must therefore have been predominantly North–

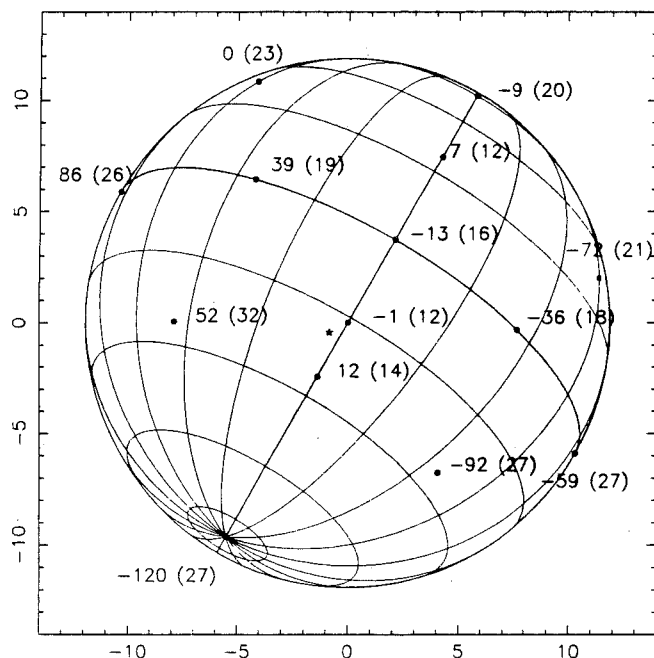


FIG. 3.—Line-of-sight beam-averaged winds (m s^{-1}) at $39\text{--}63 \text{ km}$ (measured between 0 and 1 MHz from line center). The Doppler shifts have been corrected for the solid rotation of Mars. Positive winds indicate recession. The number in brackets indicates the 1σ uncertainty. The star locates the subsolar point.

TABLE 2
EQUATORIAL VELOCITIES FOR THE ZONAL FLOW MODELS

Case Number	Fitting Interval (MHz)	Latitude Range of Zonal Flow	Best-Fit Model Space $V_{eq}(m s^{-1})^a$	Minimum χ^2 Model $V_{eq}(m s^{-1})$
1.....	0-1	$-90^\circ \leq \beta \leq 90^\circ$	55 to 165	110
2.....	0-1	$-50 \leq \beta \leq 0$	85 to 290	185
3.....	0-1	$0 \leq \beta \leq 50$ ^b ^b
4.....	0-1	$-20 \leq \beta \leq 20$	100 to 230	160
5 ^c	0-1	$-60 \leq \beta \leq 20$	80 to 240	160
6.....	0-0.35	$-90 \leq \beta \leq 90$	115 to 170	145
7.....	0.35-1	$-90 \leq \beta \leq 90$	105 to 195	150

^a Positive velocities indicate easterlies.

^b No acceptable model space.

^c Sinusoidal rolloff centered at 20 South and zeros at 20 North and 60 South (consistent with Pollack et al. 1990).

South. However, North and South offsets were found to have no effect on the model spaces of Table 2.

Finally, model spaces were generated independently for the 0-0.35 MHz and 0.35-1 MHz data. The acceptable ranges in V_{eq} (Table 2) overlapped the 0-1 MHz data, indicating that the easterly flow exists in both the 39-56 km and 52-78 km altitude ranges, though wind strength variation with altitude could not be discerned.

Regarding the Table 2 cases, while the magnitude of the χ^2 uncertainty is high, the signature of a strong easterly flow is clearly evident, and allows coarse comparison to dynamical models.

5. DISCUSSION

Theoretical models of the Mars middle atmosphere are at a very early stage of development, owing to limited atmospheric data sufficient to provide usable constraints. However, the detection of polar warming at 50-85 km altitude (Deming et al. 1986; Rothermel et al. 1988) prompted the development of simple dynamical models to estimate the likely role of gravity wave drag on the circulation of the Mars middle atmosphere (Jaquin 1989; Barnes 1990). Also, Mars GCM model fields are now being calculated up to nearly 50 km, thereby providing a crude estimate of middle atmosphere zonal winds (Pollack et al. 1990).

The thesis work of Jaquin (1989) adapted a two-dimensional, linearized, axisymmetric formulation of the equations of motion (Leovy 1964) to Mars. His code made use of calculated radiative equilibrium temperatures, Rayleigh friction and Newtonian cooling time scales (as a function of log pressure) to drive the Mars model circulation. From simulation of the growth and breakdown of gravity waves over realistic Martian terrains, the wave breaking level, on a global average basis, was set at 50 km, and the derived Rayleigh friction time constant above 50 km was 1 day, much less than the Newtonian cooling time constant. The velocity fields obtained indicate that the Mars middle atmosphere is characterized by a pole-to-pole circulation at the solstices, and an equator-to-pole circulation at the equinoxes. As in the terrestrial mesosphere, the Coriolis torques exerted on this meridional flow generate mean zonal easterlies (retrograde) in the summer hemisphere, and westerlies (prograde) in the winter hemisphere. Zonal wind speeds in the core of the winter zonal jet, being nearly 10 times faster than the meridional winds, are on the order of $80 m s^{-1}$, with the jet centered near 60° latitude and about 45 km. The

summer easterly zonal jet is weaker, broader in latitudinal extent, and its core is closer to the equator.

The Barnes (1990) model formulation is somewhat more realistic in that the wave breaking level (where wave drag effects maximize) is allowed to vary with season. His results indicate that the Deming et al. (1986) picture of the Mars thermal structure in the 50-85 km region, with both poles warm and the tropics cold, corresponds to a relatively low-level (50 km) winter westerly jet and a relatively high (85 km) summer easterly jet. In the winter hemisphere, plausible mesoscale gravity waves, generated by flow over topography, seem able to force the closure of the westerly jet in the 40-80 km region. The corresponding meridional circulation acts to produce very warm winter polar temperatures at these altitudes with values 50 K above radiative equilibrium, comparable to those seen by Deming et al. (1986). The core of the winter westerly jet (channel center at 60° latitude) is obtained at 50 km, with a wind speed of about $150 m s^{-1}$. The zonal wind decreases very sharply above this level to values of about $25 m s^{-1}$ at 80-90 km, where meridional winds become appreciable.

In the summer hemisphere, "filtering" by low atmosphere (below ~ 15 km) westerlies is especially effective in keeping stationary waves (forced by flow over topography) from propagating into the middle atmosphere. Results indicate that gravity waves might force the closure of the summer (easterly) jet at higher levels than in the winter hemisphere. In the 50-80 km region, summer high latitudes might be warm, as indicated by Deming et al. (1986). The summer pole would be cold only at still higher altitudes due to adiabatic cooling. Nominal summer zonal winds peak at 85 km, with easterly wind speeds of $130 m s^{-1}$; winds at 50 km are on the order of $100 m s^{-1}$, somewhat weaker than the corresponding winter hemisphere westerlies.

Both of these models specifically address the Mars middle atmosphere dynamics. However, they are based upon the simplified equations describing the winds. Lower atmosphere circulation models are available that are based upon the full-blown primitive equations and Mars radiative, chemical, and dust formulations. These models include a general circulation model—GCM (Pollack et al. 1990) that calculates three-dimensional fields over the globe on a rather coarse grid from the ground up to 47 km. A higher resolution Mars climate model is under development, using somewhat elementary infrared physics, simulating winds up to 55 km (Barnes 1991, private communication). Both of these models predict relatively weak summer zonal easterlies (compared to the

westerlies), with $\sim 120 \text{ m s}^{-1}$ maximum velocities near 50 km. These jets do not appear to be "closing off" at the model upper boundaries where a "sponge layer" is assumed for numerical stability. However, the basic structure of the easterly flow pattern is a summer jet with a core near 20–25 South latitude and easterlies extending over a broad latitude range from 60 South to 20 North latitude. The winds measured here at 40–70 km, when characterized by a sinusoidal latitudinal distribution centered at 20 South (60 South to 20 North), yield a core speed of $160 \pm 80 \text{ m/s}$, generally consistent with the Pollack et al. (1990) GCM simulations.

Due to the poor coverage of the Northern hemisphere, our observations cannot firmly establish the direction nor the strength of the wind there. Specifically, models with and without winds in the Northern hemisphere fit the data equally well. Therefore theoretical models, predicting the intrusion of summer easterlies into the winter hemisphere (as is common in the terrestrial middle atmosphere; Holton 1975) and the closure of the winter westerly jet remain to be observationally tested.

On the whole, these wind measurements combined with the temperature measurements of Deming et al. (1986) appear consistent with current theoretical models. The CO millimeter-wave measurements also provide the mapping of the atmospheric temperature. For a coherent picture of the middle atmosphere dynamics, the winds inferred here at 40–70 km should be viewed in the context of derived latitudinal tem-

perature variations. This unfortunately does not give convincing results because (1) in the absence of absolute calibration the retrieved temperatures are uncertain by $\pm 20 \text{ K}$ (Lellouch et al. 1991) and (2) the large size of the beam considerably smears out any temperature latitudinal contrast, especially in the Northern hemisphere.

Finally, the beam-integrated wind velocity at point 4, inconsistent with a zonal flow, can be interpreted as a high-velocity meridional component restricted to the South polar region. The beam-integrated blueshift is consistent with equatorial-ward transport. Meridional velocities of this magnitude ($\geq 100 \text{ m s}^{-1}$) seem high, yet need to be validated using theoretical models for Mars conditions of this particular time period.

Future observational progress on the middle atmosphere of Mars dynamics requires simultaneous measurement of the temperature and wind velocities. Infrared heterodyne measurements at $10 \mu\text{m}$ obtained during the 1988, 1990 oppositions, admitting tenfold better spatial resolution than the microwave, are ideally suited for this study (Goldstein et al. 1991, in preparation). Additionally, existing GCMs will have to be extended to significantly higher altitudes.

We are indebted to Fred Espenak for kindly computing the Martian ephemeris. We also thank Jeff Barnes and Jim Pollack for illuminating discussions. J. J. G. was supported by a grant from the Smithsonian Institution Scholarly Studies Program.

REFERENCES

- Barnes, J. R. 1990, *J. Geophys. Res.*, 95, 1401
 Buhl, D., Goldstein, J. J., & Chin, G. 1991, *ApJ*, submitted
 Deming, D., et al. 1986, *Icarus*, 66, 366
 Goldstein, J. J. 1990, NASA CR 4290, Absolute Wind Measurements in the Lower Thermosphere of Venus using Infrared Heterodyne Spectroscopy, (Washington, DC: GPO)
 Goldstein, J. J., et al. 1991, *Icarus*, submitted
 Haberle, R. M., Leovy, C. B., & Pollack, J. B., 1982, *Icarus*, 50, 322
 Hanel, R., et al. 1972, *Icarus*, 17, 423
 Hess, S. L., et al. 1977, *J. Geophys. Res.*, 82, 4559
 Holton, J. R. 1975, *Meteorological Monographs*, Vol. 15, No. 37, The Dynamic Meteorology of the Stratosphere and Mesosphere (Boston: Amer. Meteorological Society)
 Holton, J. R. 1982, *J. Atmos. Sci.*, 39, 791
 ———. 1983, *J. Atmos. Sci.*, 40, 2497
 Holton, J. R., & Zhu, X. 1984, *J. Atmos. Sci.*, 41, 2653
 Jaquin, R. F. 1989, Ph.D. thesis, Cornell University
 Lellouch, E., Paubert, G., & Encarnaz, T. 1991, *Planet Space Sci.*, 39, 219
 Leovy, C. B., 1964, *J. Atmos. Sci.*, 21, 327
 Lindzen, R. S. 1981, *J. Geophys. Res.*, 86, 9707
 Martin, T. Z. 1981, *Icarus*, 45, 427
 Pierce, K., Muhleman, D. O., & Berge, G. L. 1991, *Icarus*, in press
 Pollack, J. B., Haberle, R. M., Schaeffer, J., & Lee, H. 1990, *J. Geophys. Res.*, 95, 1447
 Rothmel, H., Kaufl, H. U., Schrey, U., & Drapatz, S. 1988, *Astron. Astrophys.*, 196, 296

Article VI

OBSERVATIONS OF PARENT MOLECULES IN COMETS P/BORSSEN-METCALF (1989o), AUSTIN (1989c1) AND LEVY (1990c) AT MILLIMETRE WAVELENGTHS: HCN, H₂S, H₂CO AND CH₃OH

D. Bockelée-Morvan¹, J. Crovisier¹, P. Colom¹, D. Despois² and G. Paubert³

¹ Observatoire de Paris, Section de Meudon, F-92195 Meudon, France

² Observatoire de Bordeaux, BP 89, Avenue Pierre Sémirot, F-33270, Floirac, France

³ IRAM, Avenida Divina Pastora, 7, N.C., E-18012 Granada, Spain

ABSTRACT

Because of their small size and their long-term cold storage in the distant Oort cloud, comets are thought to be the more primitive among Solar System bodies, and fully representative of the composition of the primordial solar nebula. One of the privileged tools to identify and study the parent molecules directly sublimed from cometary nuclei is microwave spectroscopy. We report on millimetre observations conducted with the IRAM 30-m telescope in comets P/Brorsen-Metcalf (1989o) on September 1989, Austin (1989c1) on April and May 1990, and Levy (1990c) on August 1990. Two species were detected for the first time in a comet: hydrogen sulfide H₂S and methanol CH₃OH. In addition, we detected rotational transitions of HCN and H₂CO. Preliminary estimates of molecular abundances are presented. Searches for SO₂, OCS, H₂CS and HC₃N, although negative, give significant constraints on the abundances of these molecules. These observations may help to resolve whether comets formed from unaltered interstellar material.

Keywords: Comets, molecules, millimetre lines

1. OBSERVATIONS

1.1 Description

The observations were performed with the 30-m telescope of the Institut de Radio Astronomie Millimétrique (IRAM) at Pico Veleta in Spain. Comet P/Brorsen-Metcalf (1989o) was observed on September 2-7, 1989 at $r_h = 0.51$ AU from the Sun and $\Delta = 1$ AU from the Earth (Ref. 1). Comet Austin (1989c1) was first observed on April 2, 1990 (close to its perihelion) at $r_h = 0.41$ AU and $\Delta = 1.19$ AU, and secondly in the May 21-25, 1990 period (around its perigee) at $r_h = 1.1$ AU and $\Delta = 0.24$ AU. Comet Levy (1990c) was observed on August 26-31, 1990 at $r_h = 1.4$ AU and $\Delta = 0.43$ AU, around its perigee. We used three SIS mixer receivers simultaneously (83-115 GHz, 130-170 GHz, 209-270 GHz), which allowed us to search for different lines at the same time. The spectrometers consisted in two banks of 128x100 kHz channels (corresponding to a velocity resolution of 0.34 km s⁻¹ and 0.11 km s⁻¹ for the HCN lines at 89 GHz and 266 GHz respectively), two banks of 512x1 MHz channels and an Acousto Optical Spectrometer (AOS) with a 500 MHz bandwidth and a resolution of 1.2 MHz. In order to study the extension of the coma, observations of comet Austin at 6 arcsec and 12 arcsec North, East, South and West from the nucleus were performed during the May observing period. The observed integrated line areas, or 3-sigma upper limits, are listed in Table 1.

1.2 Hydrogen cyanide

The J(1-0) (89 GHz) and J(3-2) (266 GHz) rotational transitions of HCN were marginally detected in comet P/Brorsen-Metcalf. The HCN J(1-0) transition was only detected previously in comet Kohoutek 1973 XII and P/Halley despite many searches in other comets (Ref. 2). Comet P/Brorsen-Metcalf is the first comet in which the HCN J(3-2) transition has been detected.

On April 2, the HCN J(1-0) line was not detected in comet Austin. However, during the May observing session, both HCN J(1-0) and J(3-2) transitions were firmly observed in the comet (Refs. 3, 4): some selected integrated spectra are shown in Figs. 1 and 2.

The HCN J(1-0) and J(3-2) transitions were detected in comet Levy, both with a signal-to-noise ratio exceeding 10. The J(1-0) line gave the strongest radio signal ever observed in a comet, with a brightness temperature integrated over the line close to 4 K km s⁻¹ (Ref. 5).

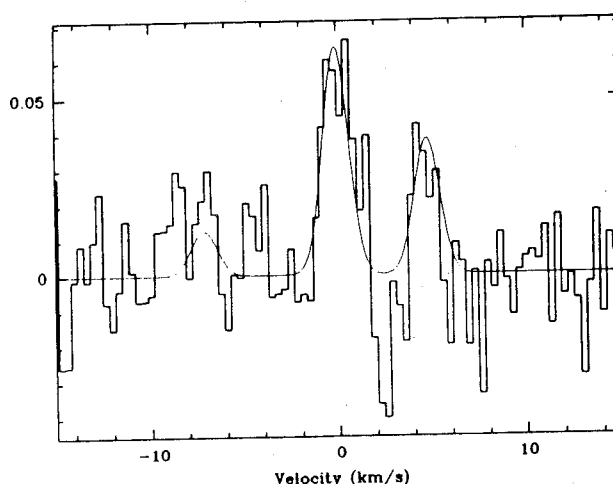


Figure 1. The HCN J(1-0) spectrum of comet Austin observed on May 21 to 23, 1990 (average of centre and offset positions). The result of a Gaussian fitting of the three HCN hyperfine components is shown. The intensity scale is main-beam antenna brightness temperature in K. The velocity scale is with respect to the comet rest velocity, for the F 2-1 main component.

1.3 Formaldehyde

The observations of the H_2CO $3_{12}-2_{11}$ transition at 226 GHz in comet P/Brorsen-Metcalf lead only to a marginal detection, with a signal-to-noise ratio of 4. However, it gave the first conclusive evidence for the possibility of observing cometary formaldehyde at millimetre wavelengths. Formaldehyde was not detected in comet Austin in April, but was easily detected in May (Ref. 3), with a S/N ratio larger than 10 for the average spectrum obtained for May 21, 22, 25 shown in Fig. 3. Formaldehyde at 226 GHz was also detected in comet Levy (Ref. 5), with a S/N ratio on the order of 6 (Table 1).

1.4 Hydrogen sulfide

Hydrogen sulfide H_2S was unambiguously detected in comet Austin on May 23, 24 and 25, 1990, through its $1_{10}-1_{01}$ ortho transition at 169 GHz (Ref. 4), with a S/N ratio of 7 for the average spectrum of May 24-25. This was the first time that H_2S was observed in a comet, although it was highly suspected from the tentative detection of H_2S^+ in the visible spectrum of comet Iras-Araki-Alcock 1983 VII (Ref. 6) and from a signal at mass 35 attributed to H_3S^+ in the Neutral Mass Spectrometer Giotto investigations of comet P/Halley (Ref. 7). In addition to the 169 GHz line (Fig. 4), the $2_{20}-2_{11}$ para transition at 217 GHz of H_2S (Fig. 5) was detected in comet Levy (Table 1, Ref. 5).

1.5 Methanol

Methanol was observed in comet Austin on May 25, 1990 through its $\text{J}(2-1) \Delta K = 0$ transitions at 97 GHz, and its $\text{J}(3-2) \Delta K = 0$ transitions at 145 GHz, using 512x1 MHz filter banks. Two $\text{J}(2-1)$ transitions are marginally present in the 97 GHz spectrum while four $\text{J}(3-2)$ lines are detected with signal to noise ratios between 2 and 5 in the 145 GHz spectrum insuring an unambiguous detection of methanol in that comet and giving the first evidence of the presence of this molecule in comets (Ref. 8). During the observing run of comet Levy, methanol was investigated at 97, 145, 165 and 219 GHz using the 1MHz and 100 kHz filter banks and the AOS: a dozen transitions were detected (Ref. 5). Fig. 6 shows the 145 GHz spectrum observed on August 27, where six lines are observed with S/N ratios between 3 and 8.

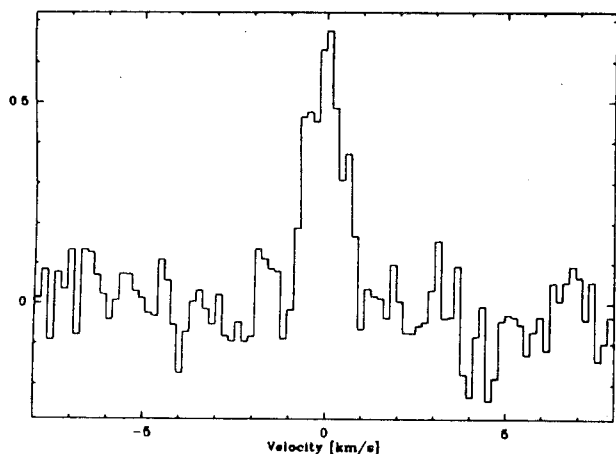


Figure 4. The H_2S $1_{10}-1_{01}$ spectrum of comet Levy observed on August 31, 1990. The intensity scale is main-beam antenna brightness temperature in K.

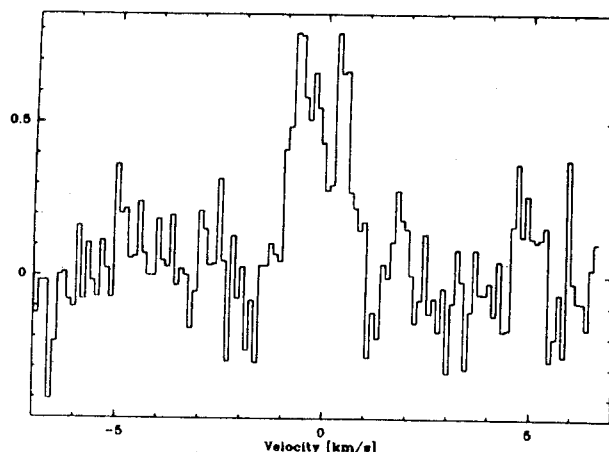


Figure 2. The HCN $\text{J}(3-2)$ spectrum of comet Austin of May 23-24, 1990 (average of centre and offset positions). The intensity scale is main-beam antenna brightness temperature in K.

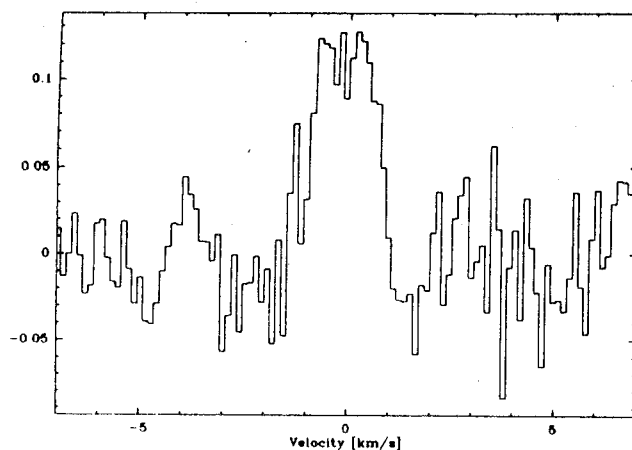


Figure 3. H_2CO $3_{12}-2_{11}$ spectrum of comet Austin observed on May 21-25, 1990 (average of centre and offset positions). The intensity scale is main-beam antenna brightness temperature in K.

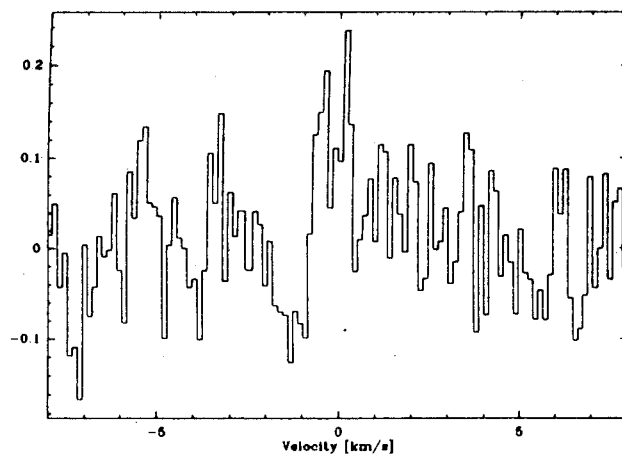


Figure 5. The H_2S $2_{20}-2_{11}$ spectrum of comet Levy observed on August 31, 1990. The intensity scale is main-beam antenna brightness temperature in K.

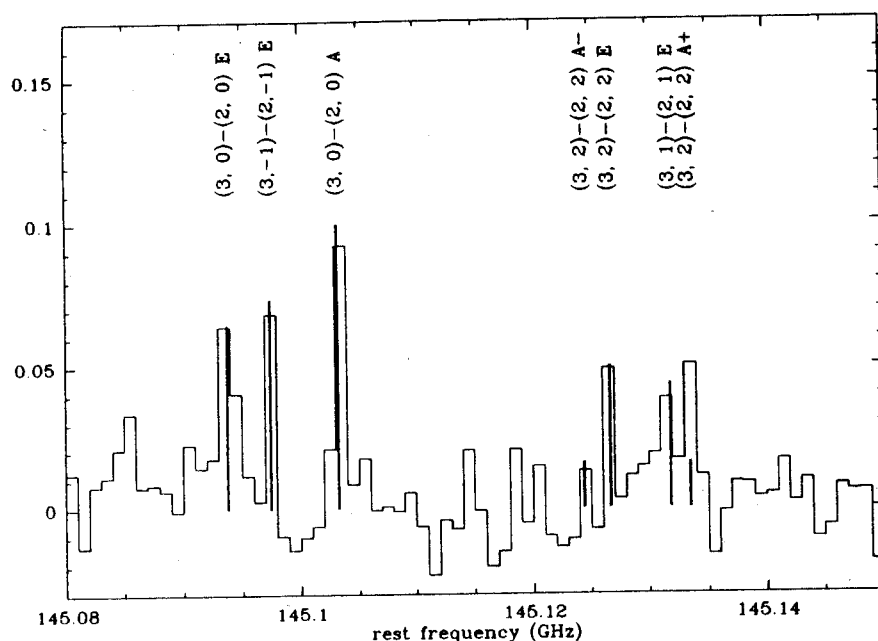


Figure 6. $J(3-2) \Delta K = 0$ transitions of CH_3OH around 145 GHz observed towards comet Levy on August 27, 1990. Vertical lines indicate the line positions and the relative intensities expected for rotational LTE at $T_{\text{rot}} = 30$ K. The intensity scale is main-beam antenna brightness temperature in K.

1.6 Other molecules

During the observing run of comet Levy, we searched for the $J(24-23)$ line of HC_3N at 218 GHz and the $J(18-17)$ line of OCS at 219 GHz. We also tried to detect the $5_{15-4_{04}}$ (135 GHz), $5_{24-5_{15}}$ (165 GHz) and $7_{17-6_{06}}$ (165 GHz) transitions of SO_2 , and the $4_{14-3_{13}}$ transition at 135 GHz of H_2CS . We got no detections. The 3-sigma upper limits are given in Table 1. Many other molecular transitions which were serendipitously observed with the AOS and 1 MHz filter banks during our observations are currently under study.

2. RESULTS

2.1 Deriving production rates

In order to convert the observed line intensities into molecular production rates or abundances with respect to water, one needs excitation models and assumptions on the molecular density distribution. It has been shown that, for the molecules investigated in the present work, the prevailing excitation mechanisms are thermal excitation by collisions and radiative excitation of the vibrational bands by solar radiation (Ref. 9). Comprehensive models treating the evolution of the excitation conditions from the collision-dominated region (inner coma) to the radiation-dominated region (outer coma) have been developed for linear molecules (such as HCN , OCS and HC_3N , Ref. 10) and formaldehyde (Ref. 11), but have not been yet achieved for the other molecules involved in this paper. Fortunately, in the case of the observations around the perigee of comets Austin and Levy, the instrumental beam was sampling mostly the collisional region of the coma: the field of view radius was on the order of 3000 (6000) km at 145 GHz and 2000 (4000) km at 225 GHz for comet Austin (Levy), while the size of the collisional region should not be much smaller than 3000 (5000) km. Therefore, the LTE model for these molecules should be a good approximation. On the other hand, whatever the model used, our evaluations heavily rely on the assumed rotational temperatures. The only exception concerns the observations of HCN and H_2CO in comet P/Brorsen-Metcalf, for which the collision region is diluted in the beam.

The production rate evaluations also depend on the adopted molecular density distribution. We assume here an isotropic and steady production from the nucleus and gas expansion at the velocity of 0.8 km s^{-1} (which is in agreement with the observed linewidths).

A problem arises, however, for formaldehyde. It is not well known, at the present time, whether H_2CO is a genuine parent molecule, or comes from a distributed source in the coma (Refs. 12, 13) which could be dust grains, polyoxymethylene (Ref. 14), or other molecules. That formaldehyde originates from a distributed source is also suggested by our observations of the 226 GHz H_2CO line in comet Austin performed at offset positions. Indeed, formaldehyde is a short-lived molecule, with a photodissociation rate of $2.8 \cdot 10^{-4} \text{ s}^{-1}$ at 1 AU from the Sun, according to Ref. 15. The line intensities expected for the 226 GHz transition at half-beam and one beam off the nucleus (6 arcsec and 12 arcsec corresponding to about 1100 km and 2200 km) are about 0.7 and 0.3 the intensity observed at the centre position, if one assumes that H_2CO comes directly from the nucleus, whereas the observed ratios are 0.8 ± 0.2 and 0.64 ± 0.14 . The density distribution of H_2CO being ill-known, it will be difficult to give accurate estimates of its abundance. We will tentatively compute production rates for both a parent molecule distribution and a daughter Haser density distribution, with a parent scalelength of 10^4 km at 0.9 AU from the Sun and scaled proportionally to r_h^2 . This density distribution copes approximately with the spatial distribution of CO measured by NMS on Giotto, interpreted in Ref. 16 as to be due to an extended source of CO in the coma.

Another fundamental parameter entering in the description of the density distribution is the molecular lifetime against photodestruction by the Solar radiation field. For HCN , we adopt the value of 40 000 s (at $r_h = 1 \text{ AU}$ from the Sun) pertaining to active Sun, as estimated in Ref. 1. From the H_2S laboratory ultraviolet absorption spectrum of Ref. 17 and the solar reference spectrum of Ref. 15, we evaluate the H_2S lifetime to 4000 s at 1 AU. We estimate the lifetime of methanol to 77000 s, from the laboratory ultraviolet absorption spectrum of Ref. 18. Lifetimes were scaled according to the r_h^2 law.

Table 1: Line intensities and production rates

Molecule	Date	$\int T_B dv^a$ [K km s ⁻¹]	Q [s ⁻¹]	Q/Q[H ₂ O]
<i>P/Brorsen-Metcalf (1989o)</i>				
HCN J(1-0)	89/09/04-07	0.18 ± 0.03	4.5E26	1.8E-3 ^b
HCN J(3-2)	89/09/02-03	0.96 ± 0.21		
H ₂ CO 3 ₁₂ -2 ₁₁	89/09/04-07	0.16 ± 0.04	1.0E27 ^c 6.8E26 ^d	4.0E-3 ^{c,b} 2.7E-3 ^{d,b}
<i>Austin (1989cl)</i>				
HCN J(1-0)	90/04/02	< 0.10	< 4.0E26	< 2.0E-3 ^b
HCN J(1-0)	90/05/23	0.20 ± 0.07	2.0E25	3.3E-4 ^e
HCN J(3-2)	90/05/23	1.27 ± 0.18		
H ₂ CO 3 ₁₂ -2 ₁₁	90/04/02	< 0.07	< 9.1E26 ^c < 6.7E26 ^d	< 4.6E-3 ^{c,b} < 3.3E-3 ^{d,b}
H ₂ CO 3 ₁₂ -2 ₁₁	90/05/21-25	0.25 ± 0.03	3.6E26 ^c 3.3E25 ^d	6.0E-3 ^{c,e} 5.5E-4 ^{d,e}
H ₂ S 1 ₁₀ -1 ₀₁	90/05/24-25	0.28 ± 0.05	1.3E26	2.1E-3 ^e
CH ₃ OH (3,0)-(2,0)A	90/05/25	0.25 ± 0.05	4.9E26	8.2E-3 ^e
<i>Levy (1990c)</i>				
HCN J(1-0)	90/08/29	0.40 ± 0.03	6.6E25	2.6E-4 ^b
HCN J(3-2)	90/08/29	3.87 ± 0.30		
H ₂ CO 3 ₁₂ -2 ₁₁	90/08/30	0.39 ± 0.06	8.3E26 ^c 9.2E25 ^d	3.3E-3 ^{c,b} 3.7E-4 ^{d,b}
H ₂ S 1 ₁₀ -1 ₀₁	90/08/31	0.77 ± 0.06	6.7E26	2.6E-3 ^b
H ₂ S 2 ₂₀ -2 ₁₁	90/08/31	0.16 ± 0.03		
CH ₃ OH (3,0)-(2,0)A	90/08/27	0.32 ± 0.04	1.8E27	7.2E-3 ^b
HC ₃ N J(24-23)	90/08/27	< 0.09	< 1.2E25	< 5.0E-5 ^b
SO ₂ 7 ₁₇ -6 ₀₆	90/08/29	< 0.17	< 6.1E26	< 2.4E-3 ^b
OCS J(18-17)	90/08/28	< 0.16	< 5.1E26	< 2.0E-3 ^b
H ₂ CS 4 ₁₄ -3 ₁₃	90/08/28	< 0.11	< 2.4E26	< 9.6E-4 ^b

a. Observations aimed at the nucleus.

b. Q[H₂O] from OH 18 cm observations: 2.5 E29 s⁻¹ for P/B-M (Ref. 32) and Levy, 2.0 E29 for Austin in April (Ref. 32).

c. Assuming a daughter Haser density distribution with a parent scalelength of 10⁴ km at 0.9 AU from the Sun and scaled as r_h^{-2} .

d. Assuming a parent distribution.

e. Q[H₂O] measured from OH IUE observations: 6.0 E28 s⁻¹ for Austin in May (Ref. 33).

2.2 Rotational temperatures

We have several indications that the rotational temperatures are low and below 50 K. As shown in Fig. 6, the transitions of methanol observed at 145 GHz in comet Levy are well explained by a rotational temperature of 30 K. The relative intensities of 145 GHz lines of methanol observed in comet Austin indicate a temperature around 20 K (Ref. 19). From the intensity ratio of the 169 and 217 GHz lines of H₂S observed simultaneously in comet Levy, we derive an excitation temperature of 40 K. Since the molecules sampled by the beam may have partly relaxed from thermal equilibrium, it is still premature to relate measured rotational temperatures to the kinetic temperature of the inner coma. Some deviations from pure LTE are indeed suggested on the 145 GHz rotational spectrum of methanol (Fig 6).

In the derivation of the HCN and H₂CO production rates, we will assume a kinetic temperature of 50 K. Rotational LTE at 50 K and 40 K is adopted for the derivation of the H₂S production rate in comet Austin and Levy respectively. For methanol, we use the excitation temperatures found in relative agreement with the rotational spectrum at 145 GHz and the observed intensity of the (3,0)-(2,0) A line at 145.103 GHz (Table 1). 3-sigma upper limits on the abundances of OCS and HC₃N are computed using Ref. 10

and a kinetic temperature of 50 K. For H₂CS and SO₂, we assume LTE at 50 K.

2.3 Molecular abundances

The molecular production rates and abundances relative to water computed according to the preceding sub-sections are reported in Table 1.

The HCN abundance measured in P/Brorsen-Metcalf (Table 1) is roughly in agreement with that measured in P/Halley ($\sim 10^{-3}$), but in comets Austin and Levy it is smaller by about a factor of 3. A similar deficiency was found in comet Wilson (1987 VII) (Ref. 20), suggesting a chemical difference in periodic and new comets. The HCN production rates appear to be significantly smaller than the CN production rates measured by optical observations (Refs. 21, 22), indicating that another source for CN radicals might be present in comets. This is in agreement with previous observations (Ref. 2). The upper limit obtained on the abundance of HC₃N ($[HC_3N]/[H_2O] < 5 \cdot 10^{-5}$) shows that this molecule is much less abundant than HCN in comets. This is consistent with $[HC_3N]/[HCN]$ abundance ratios found in interstellar clouds (Ref. 23). HC₃N is a possible parent molecule for CN, but is not the major lacking source of CN radicals.

The derived H_2CO relative abundances are on the order of 0.5 %, an order of magnitude smaller than those estimated in P/Halley. Combes et al. (Ref. 24) derived a relative production rate of 0.04 from the IKS-Vega experiment, but they state that this may be an over-estimate due to the contamination of the formaldehyde bands by other species. Snyder et al. (Ref. 12) infer 0.015 from their VLA observations of the 6 cm line, but the examination of the excitation conditions of this line (Ref. 11) casts some doubt on the reliability of this detection. We can note that our measured abundances are still much smaller when we assume that H_2CO escapes from the nucleus (Table 1).

H_2S is a minor component of cometary nuclei, with an abundance relative to water on the order $2 \cdot 10^{-3}$. In their analysis of the Giotto NMS data, Marconi et al. (Ref. 7) obtained an unreasonably large H_2S production rate when assuming a parent molecule distribution for this molecule. In fact, this problem arises because they used a too short lifetime against photodissociation. When using a lifetime of 4000 s (Sect. 2.1), the NMS data become consistent with a relative abundance similar to that we observed in comet Austin and Levy. $[\text{H}_2\text{S}]/[\text{H}_2\text{O}]$ is an order of magnitude less than the cosmic relative abundance $[\text{S}]/[\text{O}] = 0.02$. The total sulfur abundance of cometary volatiles $[\text{S}]/[\text{H}_2\text{O}]$ is also estimated to be on the order of 0.02 (Ref. 25). The only other sulfur species firmly observed in comets are S_2 and CS (Ref. 26), with relative abundances still smaller than of H_2S . The significant upper limits obtained for OCS, SO_2 and H_2CS indicate that the lacking sulfur is not contained in these species. Thus cometary sulfur might be retained in still undetected molecules.

As a preliminary estimate, the methanol abundance relative to water is found to be on the order of 10^{-2} in both comets Austin and Levy.

3. CONSEQUENCES FOR COMETARY FORMATION

Methanol is known to be an abundant constituent of the interstellar medium: its abundance relative to water in the solid phase was recently evaluated to 0.07 in the protostellar source W33A (Ref. 27). Methanol cannot survive at high temperatures or in the presence of a UV field. It has not been yet identified in other solar-system bodies. Therefore, the existence of cometary methanol with a significant abundance is a strong evidence that cometary nuclei retained, at least partly, some of the original matter of the presolar nebula.

The case for hydrogen sulfide is somewhat different, since its abundance in the interstellar gas is low, either from observations (Ref. 28) or from model predictions of gas-phase interstellar chemistry (Ref. 29). It must be noted, however, that H_2S has been detected through its infrared signature in the solid phase in W33A (Ref. 30), with a yet unknown abundance. Before concluding that this molecule is overabundant in cometary matter compared to interstellar matter, the abundance of H_2S trapped in interstellar ices and the role of grain-surface reactions should be assessed.

Since these parent molecules can only condense at low temperatures, their presence in cometary nuclei give severe constraints on the physical conditions of cometary formation. Yamamoto (Ref. 31) has computed that, for a gas density of 10^{13} molec. cm^{-3} (which is thought to be typical of the presolar nebula) CH_3OH and HCN can only condense at temperatures lower than 99 and 95 K, respectively. This is to be compared to 152 K, for water. For H_2CO and H_2S , these temperatures are still lower: 64 and 57 K, respectively. For CO, the condensation temperature

is as low as 25 K, but the presence of CO as a genuine cometary parent molecule is still controversial. Thus, the detection of cometary H_2S is a further evidence that comets were formed and maintained in cold conditions.

4. CONCLUSION

In addition to the ubiquitous OH radical, four molecular species have now been unambiguously identified in cometary atmospheres by radio spectroscopy. This technique proved to be able to study relatively complex organics such as methanol, or relatively minor species such as HCN or H_2S . It promises to be a productive mean of determining the composition of comets. Therefore, microwave spectroscopy offers us the possibility of reconstructing the conditions under which comets formed and may help in tracing the history of the Solar System.

5. REFERENCES

- Colom, P., Crovisier, J., Bockelée-Morvan, D., Despois, D. and Paubert, G. 1990, *Astr. Ap.* (submitted)
- Crovisier, J. and Schloerb, F.P. 1990, in *Comets in the Post-Halley Era*, Eds. R. Newburn and J. Rahe, Kluwer Academic Publishers, Dordrecht (in press)
- Bockelée-Morvan, D., Despois, D., Paubert, G., Colom, P. and Crovisier, J. 1990, *IAU Circ.* No 5020
- Crovisier, J., Despois, D., Bockelée-Morvan, D., Paubert, G. and Colom, P. 1990, *IAU Circ.* No 5022
- Colom, P., Crovisier, J., Paubert, G., Bockelée-Morvan, D. and Despois, D. 1990, *IAU Circ.* No 5087
- Cosmovici, C.B. and Ortolani, S. 1984, *Nature* 310, 122
- Marconi, M.L., Mendis, D.A., Korth, A., Pin, R.P., Mitchell, D.L., and Reme, H. 1990, *Ap. J.* 352, L17
- Despois, D., Bockelée-Morvan, D., Paubert, G., Colom, P. and Crovisier, J. 1990, *IAU Circ.* No 5027
- Crovisier, J. and Encrenaz, E. 1983, *Astr. Ap.* 126, 170
- Crovisier, J. 1987, *Astr. Ap. Suppl. Ser.* 68, 223
- Bockelée-Morvan D. and Crovisier, J. 1990, (in preparation)
- Snyder, L.E., Palmer, P. and de Pater, I. 1989, *Ap. J.* 97, 246
- Krankowsky, D. 1990, in *Comets in the Post-Halley Era*, Edts R. Newburn and J. Rahe, Kluwer Academic Publishers, Dordrecht (in press)
- Boice, D.C., Huebner, W.F., Sablik, M.J. and Konno, I. 1990, *Geophys. Res. Letts* (submitted)
- Huebner, W.F. and Carpenter, C.W. 1979 *Los Alamos Informal Report* LA-8085-MS
- Eberhardt, P. et al. 1987 *Astr. Ap.* 187, 481
- Lee, L.C., Wang, X. and Suto, M. 1987, *J. Chem. Phys.* 86, 4353

18. Nee, J.B., Suto, M. and Lee, L.C. 1985, *Chem. Phys.* 98, 147
19. Colom, P., Despois, D., Bockelée-Morvan, D., Crovisier, J. and Paubert, G. 1990, in *Observations of Recent Comets*, ed. Huebner W.F., (in press)
20. Crovisier, J., Despois, D., Bockelée-Morvan, D., Gérard, E., Paubert, G., Johansson, L.E.B., Ekelund, L., Winnberg, A., Ge, W., Irvine, W.M., Kinzel, W.M. and Schloerb, F.P. 1990, *Astr. Ap.* 234, 535
21. Schleicher, D.G., Millis, R.L., Osip, D.J. and Thompson, A. 1990, in *Observations of Recent Comets*, ed. Huebner W.F., (in press)
22. Schleicher, D.G., Millis, R.L. and Osip, D.J. 1990, *IAU Circ.* No 5081
23. Irvine, W.M., Schloerb, F.P., Hjalmarson, A. and Herbst, E. 1985, in *Protostars and Planets II*, eds. D.C. Black and M.S. Matthews, The University of Arizona Press, p. 579
24. Combes et al. 1988, *Icarus* 76, 404
25. Azoulay, G. and Festou, M.C. 1985, in *Asteroids, Comets and Meteors II*, eds C.-I. Lagerkvist, B.A. Lindblad, H. Lundstedt, H. Rickman, Uppsala, Sweden
26. A'Hearn, M.F., Feldman, P.D. and Scheicher, D.G. 1983 *Ap. J.* 274, L99
27. Baas, F., Grim, R.J.A., Geballe, W., Schutte, W. and Greenberg, J.M. 1988 in *Dust in the Universe*, eds Bailey, M.E. and Williams, D.A., Cambridge University Press, p 55
28. Minh, Y.C., Irvine, W.M. and Ziurys L.M. 1989, *Ap. J.* 345, L63
29. Millar, T.J. and Herbst, E. 1990, *Astr. Ap.* 231, 466
30. Geballe, T.R., Baas, F., Greenberg, J.M. and Schutte, W. 1985, *Astr. Ap.* 146, L6
31. Yamamoto, T. 1985, *Astr. Ap.* 142, 31
32. Bockelée-Morvan, D., Crovisier, J., Gérard, E. and Bourgois, G. 1990, in *Observations of Recent Comets*, ed. Huebner W.F., (in press)
33. Budzien, S.A., Feldman, P.D., Roettger, E.E., Festou, M.C. and A'Hearn, M.F. 1990, in *Observations of Recent Comets*, ed. Huebner W.F., (in press)

